RUHR ECONOMIC PAPERS

Mark A. Andor
Manuel Frondel
Colin Vance

Mitigating Hypothetical Bias – Evidence on the Effects of Correctives from a Large Field Study

RUHR UNIVERSITÄT BOCHUM RUB rwi

#480

Imprint

Ruhr Economic Papers

Published by

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

Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
Hohenzollernstr. 1-3, 45128 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #480

Responsible Editor: Manuel Frondel

ISSN 1864-4872 (online) – ISBN 978-3-86788-544-7

Ruhr Economic Papers #480

Mark A. Andor, Manuel Frondel, and Colin Vance

Mitigating Hypothetical Bias – Evidence on the Effects of Correctives from a Large Field Study

RUHR UNIVERSITÄT BOCHUM RUB

rwi

Bibliografische Informationen der Deutschen Nationalbibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über: *http://dnb.d-nb.de* abrufbar.

http://dx.doi.org/10.4419/86788544
ISSN 1864-4872 (online)
ISBN 978-3-86788-544-7

Mark A. Andor, Manuel Frondel, and Colin Vance[1]

Mitigating Hypothetical Bias – Evidence on the Effects of Correctives from a Large Field Study

Abstract

The overestimation of willingness-to-pay (WTP) in hypothetical responses is a well-known finding in the literature. Various techniques have been proposed to remove or, at least, reduce this bias. Using responses from a panel of about 6,500 German households on their WTP for a variety of power mixes, this article undertakes an analysis that combines two common ex-ante approaches – cheap talk and consequential script – with the ex-post certainty approach to calibrating hypothetical WTP responses. Based on a switching regression model that accounts for the potential endogeneity of respondent certainty, we find that while neither the cheap-talk nor the consequential script corrective bears on the estimates of WTP, there is evidence for a lower WTP among those respondents who classify themselves as definitely certain about their answers.

JEL Classification: D12, Q21, Q41

Keywords: Willingness-to-pay; cheap talk; certainty approach

April 2014

1 Mark A. Andor, RWI; Manuel Frondel, RWI and RUB; Colin Vance, RWI and Jacobs University Bremen. *– We are grateful for invaluable comments and suggestions by Christoph M. SCHMIDT. This work has been partly supported by the Collaborative Research Center "Statistical Modeling of Nonlinear Dynamic Processes" (SFB 823) of the German Research Foundation (DFG), within the framework of Project A3, "Dynamic Technology Modeling". We also gratefully acknowledge financial support by the German Federal Ministery of Education and Research (BMBF) under grant 01LA1113A.**– All correspondence to Manuel Frondel, RWI, Hohenzollernstraße 1-3, 45128 Essen, Germany, e-mail: frondel@rwi-essen.de*

1 Introduction

Inferences on willingness-to-pay (WTP) for non-market goods ideally rely on actual, rather than hypothetical choices, that is, on revealed, rather than stated preferences. Contrary to stated preferences, revealed preferences entail a real economic commitment or consequence, such as signing a contract or the delivery of a commodity (HARRISON, 2006:125). Frequently, however, revealed-preference information is unavailable.

One reason is due to the lack of markets or third-party financing (BLUMENSCHEIN et al., 2008:114). Another reason for the absence of revealed preferences owes to market interventions that distort price signals. For example, in many European countries, contracts on the delivery of electricity produced from renewable energy technologies – so-called green electricity – are often cheaper than contracts on conventional electricity, although the cost for renewable electricity production is typically higher.[1] In Germany, for instance, this contradiction exists because for such contracts cheap green electricity, e. g. produced on the basis of competitive water power, is frequently imported from abroad, whereas it is prohibited to employ electricity based on expensive technologies, such as solar power, that are subsidized through feed-in tariffs financed by all consumers. This circumstance prevents researchers from receiving comprehensive information on consumers' true preferences, which would only be revealed by engagements in contracts on green electricity that reflect the actual cost.

In such a situation, estimating the WTP for green electricity requires methods for eliciting stated preferences that, ultimately, are based on hypothetical choices. There is ample empirical evidence, though, that hypothetical responses sometimes substantially overestimate WTP. This overestimation problem, referred to in the literature as hypothetical bias (BLUMENSCHEIN et al., 2008:115), has been documented extensively, including the laboratory experiments by CUMMINGS et al. (1995, 1997) comparing real and hypothetical WTP, as well as the meta-analysis by LIST and GALLET (2001) and

[1]The electricity production based on renewable technologies can be regarded as a public good (see e. g. MENGES et al. (2005:432), as nobody can be excluded from the associated positive effects and there is non-rivalry in the consumption of the benefits.

the reviews by HARRISON (2006) and HARRISON and RUTSTRÖM (2008).

Various techniques have been proposed to remove or, at least, reduce this bias. Three of the most prominent techniques are the consequential-script corrective, the cheap-talk protocol introduced by CUMMINGS and TAYLOR (1999) and the certainty approach. In one variant of the certainty approach, hypothetical WTP responses are divided into two classes of certainty using a follow-up question: Subjects are asked whether they are 'fairly' or 'absolutely sure' about their WTP responses, as was done by JOHANNESSON et al. (1998) in a early form of this version. Subsequently, the certainty question was adjusted by BLUMENSCHEIN et al. (1998), who only treated those hypothetical responses about which respondents were 'definitely sure' as yes-responses, whereas 'probably sure' responses were treated the same as no-responses. According to BLUMENSCHEIN et al. (1998, 2001, 2008), this approach was effective in removing hypothetical bias both in laboratory and field experiments.

While the certainty approach involves a follow-up question, the consequential and cheap-talk correctives precede the elicitation of the WTP and present respondents with scripts that are intended to encourage deeper reflection on the implications of their responses. The consequential corrective, also called consequentialism, was suggested by BULTE et al. (2005:334), but inspired by the work of CARSON et al. (2004). It is based on a script with which subjects are told that their responses to valuation questions will have real consequences. The cheap-talk approach consists of a script including an explicit discussion on the notion of hypothetical bias and its causes (see e. g. CARLSSON et al., 2005:149, WHITEHEAD, CHERRY, 2007:252), thereby asking respondents to adjust for this bias in stating their WTP.

The evidence for the impact of both approaches is inconclusive. While CUMMINGS and TAYLOR (1999) find that the cheap-talk corrective reduces bias in experimental referenda about donations to public goods, subsequent studies yield mixed results. In a second-price auction for sports cards, for example, LIST (2001) finds that this approach removed the hypothetical bias for non-dealers, yet not for dealers. Moreover, the bias was not fully removed in a study by MURPHY et al. (2005) on a voluntary contribution mechanism. In contrast, in a field study, BULTE et al. (2005) provide support for the hypothesis that stated WTP values obtained via cheap-talk and consequential treatments are lower than without inclusion of these protocols.

The present study contributes to this strand of literature with an analysis of alternative ex-ante and ex-post bias correctives in the estimation of WTP for a variety of power mixes in Germany, a country that is currently in the midst of a massive transformation of its power sector, exemplified by the legally stipulated nuclear phase-out by 2022 and the commitment to increase the share of electricity produced from renewable energy technologies from almost 25% in 2013 to 50% by 2030 and 80% by 2050. A distinguishing feature of the analysis is that we draw on almost 25,000 responses of a large panel of about 6,500 households, with which we provide for fixed-effects estimates of the WTP for distinct power mixes comprising fossil fuels, nuclear power, renewable energy, and combinations thereof, thereby controlling for unobserved heterogeneity.

Given the lack of market prices for specific power mixes, such as electricity solely produced on the basis of nuclear power, it is impossible to calibrate hypothetical WTP responses using real payments.[2] Using an experimental design, our aim is instead to gauge the extent to which the estimates of WTP vary according to two alternative ex-ante treatments in the form of cheap-talk and consequential scripts, which are crossed with an ex-post certainty procedure that endogenously divides respondents into two groups distinguished by their level of certainty.

To this end, we randomly divided the panel into three groups encompassing an equal number of households: (1) a treatment group of households who received a cheap-talk script before eliciting their WTP for five out of 14 distinct electricity mixes, (2) a second treatment group whose households received a consequential script and (3) a control group without such treatments. Upon stating their preferences, all households have been asked according to the certainty approach in the version suggested by BLUMENSCHEIN et al. (1998) whether they are probably or definitely sure

[2]Indeed, as HARRISON et al. (2004) note, it is in any case problematic to derive the true WTP for traded goods when respondents are aware of their market prices.

about their WTP responses, yielding six subgroups altogether.

Based on a switching regression model that accounts for the potential endogeneity of the certainty responses, we find that neither the cheap-talk nor the consequential script approach has a mediating effect on the estimates of WTP. Similar to BLUMENSCHEIN et al. (2008), however, we do find evidence for statistically significant differentials in the WTP according to the certainty of the respondents. Specifically, the results indicate a lower WTP among the certain respondents, a distinction that is only evident after controlling for the endogeneity of certainty status.

The subsequent section describes the survey design and the data set. Section 3 provides a description of the estimation method, followed by the presentation and interpretation of the results given in Section 4. The last section summarizes and concludes.

2 Data and Experimental Design

To elicit people's WTP for a variety of power mixes, we collaborated with the survey institute forsa, which maintains a panel of more than 10,000 households that is representative of the German-speaking population.[3] forsa collects data using a state-of-the-art tool that allows panelists to fill out the questionnaire using either a television or, if access is available, the internet. Respondents – in our survey the household heads – retrieve and return questionnaires from home and can interrupt and continue the survey at any time. A large set of socio-economic and demographic background information on all household members is available from forsa's household selection procedure and updated regularly. Within the survey period of May 10 to June 17, 2013, 6,522 households completed the questionnaire.[4]

Along the lines of LANCASTER (1966), who emphasizes that people derive utility

[3]Additional information on the panel is available at http://www.forsa.com/.

[4]A summary of the descriptive results, as well as the questionnaire, both in German, can be retrieved from the project home-page: www.rwi-essen.de/eval-map.

from both the characteristics of market goods and the consumption level, we assume that an individual's WTP for electricity specifically depends on the way it is produced. To elicit this WTP, we adopt the survey design used by GRÖSCHE and SCHRÖDER (2011), who initiated a similar survey among the forsa household panel at the outset of 2008, but did not investigate the effects of bias correctives.

The survey begins with a brief introductory text on electricity generation technologies in general. Respondents are then presented with a sequence of five randomly selected power mixes accompanied by the following text:[5] "We request that you report the maximum amount that you, personally, would be willing to pay. As a basis for comparison, please consider an electricity mix comprised exclusively of the fossil sources coal, natural gas, and oil, which has a price of €100 per month" (see Appendix A for more details).

While several formats to elicit WTP have been suggested in the literature (see Frew et al., 2003), such as the close-ended, payment scale, and bidding/bargaining formats, the open-ended format used by GRÖSCHE and SCHRÖDER (2011) has the virtue of providing a reference point while at the same time avoiding any binding restrictions on WTP bids. Responses are instead allowed to vary in a very broad range between €0 and €9,999 in discrete increments of €1. A potential drawback of the open-ended format is the possible occurrence of protest bids, wherein respondents assign either a zero or an invalidly high value to the good (HALSTEAD et al., 1992). Our empirical analysis indicates that protest bids are hardly present in our data base.

Before requesting the WTP bids for five out of a variety of 14 power mixes, panelists are randomly divided into three equally large groups, with Group 2 and 3 being either confronted with a cheap-talk or a consequential script, respectively, whereas (control) Group 1 receives no such treatments. Both scripts, presented in detail in Appendix A, are modified versions of those provided in the literature. In formulating our consequential script, we have been mainly inspired by BULTE et al. (2005:334). Follow-

[5]Randomizing the draws of the alternatives should minimize biases that may result from ordering effects (BATEMAN and LANGFORD, 1997; CLARK and FRIESEN, 2008).

ing the concise version of CARLSSON et al. (2005:149), which does not incorporate most of the characteristics found in CUMMINGS and TAYLOR (1999), our cheap-talk script is condensed as much as possible to avoid that panelists ignore the script simply because of too much time spent on reading the text.

Upon stating their preferences, all respondents are asked according to the certainty approach in the version suggested by BLUMENSCHEIN et al. (1998) whether they are probably or definitely sure about their WTP responses, yielding six subgroups altogether and a 3×2 split-sample survey design in which 2 treatments and a control group are crossed with 2 certainty levels (Table 1). The share of respondents who are definitely sure about their WTP responses, described by certainty variable C, amounts to 49.74%, implying that a slight majority of 50.26% is just 'probably sure' (Tables 1 and 2). As elaborated in Section 3, we assume that dummy variable C reflects an endogenous decision of the respondents, as opposed to their exogenous confrontation with the treatments.

Table 1: Experimental Design

	Certainty on WTP			
	Definitely Sure: $C = 1$	Probably Sure: $C = 0$	Total	Shares
Group 1 (control group)	990	1,185	2,175	33.35%
Group 2 (*cheap talk* =1)	1,180	997	2,177	33.38%
Group 3 (*consequential* =1)	1,074	1,096	2,170	33.27%
Total	3,244	3,278	6,522	100.00%
Shares	49.74%	50.26%	100.00%	

Definitions and descriptive statistics of the variables included in the model are presented in Table 2. Results for the dummy variables *cheap talk* and *consequential* indicate that non-participation in the survey did not impact the uniform distribution of the households across the two treatment groups and the control group (see also Table 1): The shares of households who belong to the cheap-talk and consequential-script groups amount to 33.4 and 33.3%, respectively. Also of note is the fact that with a share of about one third, female respondents are a minority. This is due to our decision to deliberately ask only household heads to participate in the survey, as, by definition, they typically make investment decisions and check invoices, such as electricity bills.

Table 2: Definitions of Selected Variables and Descriptive Statistics

Variable Name	Variable Definition	Mean
age	Age of respondent	52.96
female	Dummy: 1 if respondent is female	0.326
children	Dummy: 1 if respondent has children	0.208
cheap talk	Dummy: 1 if household received a cheap-talk script	0.334
consequential	Dummy: 1 if household received a consequential script	0.333
C	Dummy: 1 if household ticked the option 'definitely sure' for the certainty question	0.497
degree	Dummy: 1 if household head has a college preparatory degree	0.399
low income	Dummy: 1 if net monthly household income is lower than €1,251	0.189
medium income	Dummy: 1 if net monthly household income is between €1,251 and €2,750	0.452
high income	Dummy: 1 if net monthly household income is between €2,751 and €4,250	0.253
very high income	Dummy: 1 if net monthly household income exceeds €4,250	0.106
1-person hh	Dummy: 1 if # household members = 1	0.276
2-person hh	Dummy: 1 if # household members = 2	0.435
3-person hh	Dummy: 1 if # household members = 3	0.151
4-person hh	Dummy: 1 if # household members = 4	0.100
> 4-person hh	Dummy: 1 if # household members > 4	0.038
P	Dummy: 1 respondent has correctly indicated the broad range of average electricity prices	0.183
L	Dummy: 1 respondent has correctly indicated the broad range of the levy for renewables	0.306

The descriptive statistics for the dependent variable WTP are presented in Table 3. In interpreting the table, it bears noting that the political economy of electricity provision in Germany has been strongly influenced by two factors in recent years, both of which are reflected in the WTP figures for the alternative power mixes. The first is the country's ongoing commitment to increase the share of renewable energies, with green electricity production amounting to a share of some 25% of gross consumption by the end of 2013.

Table 3: WTP for a Variety of Electricity Mixes Relative to Electricity Production based on 100% Fossil Fuels and Comparison with the Results of GRÖSCHE and SCHRÖDER (2011).[6]

			Our Study				GRÖSCHE and SCHRÖDER (2011)			
Shares in Electricity Mix			Relative WTP 2013				Relative WTP 2008			
Fossil Fuels	Renew-ables	Nuclear Power	# of Obs.	Median	Mean	Std. Dev.	# of Obs.	Median	Mean	Std. Dev.
75%	25%	0%	2,184	100.0	96.2	60.1	1,008	100	97	29.7
50%	50%	0%	2,168	100.0	105.0	112.7	1,056	100	101	30.8
25%	75%	0%	2,099	100.0	103.3	232.0	1,031	102	106	32.9
0%	100%	0%	2,151	110.0	112.6	41.5	1,084	110	112	37.2
75%	0%	25%	2,112	80.0	84.6	190.1	1,063	85	85	30.4
50%	0%	50%	2,138	75.0	72.7	93.0	1,054	80	81	30.3
25%	0%	75%	2,171	70.0	66.6	81.8	951	80	76	33.4
0%	0%	100%	2,149	60.0	54.6	47.2	n. a.	n. a.	n. a.	n. a.
0%	25%	75%	2,143	75.0	81.9	272.1	1,088	80	81	33.8
0%	50%	50%	2,131	90.0	91.9	223.1	1,055	100	92	30.6
0%	75%	25%	2,173	100.0	95.6	55.2	1,058	100	99	34.6
50%	25%	25%	2,205	95.0	91.2	180.6	1,090	100	91	29.5
25%	50%	25%	2,145	100.0	95.4	138.3	1,048	100	96	29.5
25%	25%	50%	2,239	80.0	82.5	177.5	1,061	90	87	32.0
100%	0%	0%	6,522	100.0	100.0	0.0	2,948	100	100	0.0

The second is the nuclear catastrophe at Japan's Fukushima in 2011, which had a profound impact in exacerbating a longstanding skepticism in Germany on the merits of nuclear power and led to the legal stipulation of its phase-out in the same year. Recalling that the base category is 100% fossil fuels with a cost of electricity consumption normalized to €100, the highest mean WTP in Table 3 has a value of 112.6 and a medi-

[6]As not all respondents have provided us with five WTP bids, we end up with an unbalanced panel of 30,208 observations with about 2,150 responses for each alternative (see Table 3) and, on average, 4.6 instead of 5 bids.

an of 110, corresponding to 100% renewables. Conversely, the lowest WTP, at a mean of 54.6 and a median of 60, is seen for 100% nuclear power.

It is of interest to compare our figures with those obtained by GRÖSCHE and SCHRÖDER (2011) from the 2008 survey. While the WTP for most mixes encompassing nuclear power has changed substantially, the WTP for renewable mixes has remained relatively constant. Specifically, with €110 and €112.6, respectively, both the median and mean WTP for 100% green electricity have virtually remained the same. In contrast, with two exceptions in which the median WTP remain unchanged, all other median values for those mixes with a non-vanishing share of nuclear power are lower than those reported by GRÖSCHE and SCHRÖDER (2011). In short, while the Fukushima catastrophe did not provide for an additional push for renewable technologies in Germany, sympathies for nuclear power, as measured in terms of WTP, have shrunk.

Finally, a preliminary comparison of the effects of correctives, based on the mean WTP for electricity stemming from 100% renewables, indicates hardly any statistically significant differences at conventional significance levels (Table 4).[7] For instance, while most of the t-test statistics are not reported, the t-statistics shown in the last column of Table 4 reveal that there are no significant discrepancies across those who are definitely certain about their WTP and those who are not. Were all six subgroups to be selected perfectly randomly and, hence, were balanced with respect to both observable and unobservable factors, we would conclude that the correctives have muted effects. However, the subgroup of definitely certain individuals is not exogenously determined. As a result, multivariate methods described in the following section have to be employed that treat the certainty outcome as a choice variable.

[7] There is just one exception: Among those who are less certain on their WTP-bids, the mean WTP of respondents who received a cheap-talk script is statistically lower at a 1% significance level than for the control group.

Table 4: Mean WTP for 100% Renewables across Treatments

	Certainty on WTP		Tests on
	Definitely	Probably	Differences
	Sure: $C = 1$	Sure: $C = 0$	t-Statistics
Group 1 (control group)	113.4	115.5	0.76
Group 2 (*cheap talk* =1)	111.0	108.0	-1.09
Group 3 (*consequential* =1)	116.3	110.3	-1.58
Total	113.6	111.5	-1.17

3 Methodology

To cope with the endogeneity of certainty variable C, we apply a switching regression model with endogenous switching (see MADDALA 1983:223-228). The behavior of households is described by two regression equations that divide observations into two regimes, those who are certain about their WTP (Regime 1) and those who are uncertain (Regime 0):

$$WTP_{1i} = \boldsymbol{\beta}_1^T \cdot \mathbf{x}_{1i} + u_{1i}, \quad \text{if } C_i = 1 \quad \text{(Regime 1)}, \tag{1}$$

$$WTP_{0i} = \boldsymbol{\beta}_0^T \cdot \mathbf{x}_{0i} + u_{0i}, \quad \text{if } C_i = 0 \quad \text{(Regime 0)}. \tag{2}$$

In this equation system, WTP_{1i} and WTP_{0i} denote the households' individual WTP bids and $\mathbf{x}_{1i}$ and $\mathbf{x}_{0i}$ include their determinants, such as net household income, while $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_0$ are vectors of the associated parameters to be estimated.

C is a dummy variable indicating the certainty regime:

$$\begin{aligned} C_i = 1 \qquad & \text{if } \boldsymbol{\gamma}^T \cdot \mathbf{z}_i \geq u_i, \\ C_i = 0 \qquad & \text{otherwise}, \end{aligned} \tag{3}$$

where $\mathbf{z}_i$ includes factors, such as a good guess of electricity prices, P_i, and of the levy for subsidizing renewables, L_i, that may affect whether a household head i is either definitely sure about her WTP bids (Regime 1: $C_i = 1$) or just probably sure: (Regime 0: $C_i = 0$).

Since such guesses have been asked during the survey, in a first stage the unknown parameter vector γ can be estimated – up to a scale factor – using standard probit maximum likelihood methods, where, due to the indeterminacy of the scale factor, $Var(u_i) = 1$ can be assumed. In the endogenous switching regression model, the error term u_i is assumed to be correlated with both errors u_{1i} and u_{2i}, as there may be unobservable factors that are relevant for both the selection into either regime and WTP bids.[8]

The second stage equations to be estimated are

$$WTP_{1i} = \boldsymbol{\beta}_1^T \cdot \mathbf{x}_{1i} - \sigma_{1u} \cdot \mathrm{IVM}_{1i} + \varepsilon_{1i}, \text{ for } I_i = 1, \tag{4}$$

$$WTP_{0i} = \boldsymbol{\beta}_0^T \cdot \mathbf{x}_{0i} + \sigma_{0u} \cdot \mathrm{IVM}_{0i} + \varepsilon_{0i}, \text{ for } I_i = 0, \tag{5}$$

where ε_{1i} and ε_{0i} are new residuals with zero conditional mean and

$$\mathrm{IVM}_{1i} := \frac{\phi(\boldsymbol{\gamma}^T \cdot \mathbf{z}_i)}{\Phi(\boldsymbol{\gamma}^T \cdot \mathbf{z}_i)}, \qquad \mathrm{IVM}_{0i} := \frac{\phi(\boldsymbol{\gamma}^T \cdot \mathbf{z}_i)}{1 - \Phi(\boldsymbol{\gamma}^T \cdot \mathbf{z}_i)} \tag{6}$$

represent the two variants of the inverse Mills ratios, with $\phi(.)$ and $\Phi(.)$ denoting the density and cumulative density function of the standard normal distribution, respectively. When appended as extra regressors in the second-stage estimation, the inverse Mills ratios are controls for potential biases arising from sample selectivity: It is likely that intrinsically unobservable characteristics, such as carelessness about electricity bills, also affect WTP bids. If the estimated coefficients – σ_{1u} and σ_{0u} – are statistically significant, this is an indication of sample selectivity.

Identification of the model requires the specification of at least one variable that determines the first stage discrete outcome but not the second stage continuous outcome. We specify two such exclusion restrictions, both of which are based on the respondent's familiarity with electricity provision. The first is a dummy indicating whether

[8] All three terms are assumed to have a trivariate normal distribution, with mean vector zero and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{10} & \sigma_{1u} \\ . & \sigma_0^2 & \sigma_{0u} \\ . & . & 1 \end{bmatrix}.$$

the respondent correctly states the per-kWh price of electricity within an error margin of 3 cents, while the second is a dummy indicating whether the respondent correctly states the levied paid for renewable energy, also within an error margin of 1 cent. By law, this levy, which at the time of the survey was 5.3 cents, is included on every electricity bill. The levy thereby comprises roughly 19% of the average per-kWh price of electricity of 28.5 cents in 2013.

For the second-stage estimation, we insert the predicted values $\widehat{\text{IVM}}_{1i}$ and $\widehat{\text{IVM}}_{0i}$ using the probit estimates $\widehat{\gamma}$ of the first-stage estimation. Given that the variance of the residuals is heteroscedastic in nature (see MADDALA 1983:225), equations (4) and (5) should be estimated by weighted least squares using the Huber-White estimates of variance.

System (4) and (5) ultimately results from the fact that the conditional expectations of u_{1i} and u_{2i} generally differ from zero and are given by the standard formula for the expected value of a truncated normal distribution (see MADDALA 1983:224):

$$E(u_{1i}|u_i \leq \boldsymbol{\gamma}^T \mathbf{z}_i) = E(\sigma_{1u} u_i | u_i \leq \boldsymbol{\gamma}^T \mathbf{z}_i) = -\sigma_{1u} \cdot \frac{\phi(\boldsymbol{\gamma}^T \mathbf{z}_i)}{\Phi(\boldsymbol{\gamma}^T \mathbf{z}_i)} = -\sigma_{1u} \cdot \text{IVM}_{1i} \,, \tag{7}$$

$$E(u_{0i}|u_i > \boldsymbol{\gamma}^T \mathbf{z}_i) = E(\sigma_{0u} u_i | u_i > \boldsymbol{\gamma}^T \mathbf{z}_i) = \sigma_{0u} \cdot \frac{\phi(\boldsymbol{\gamma}^T \mathbf{z}_i)}{1 - \Phi(\boldsymbol{\gamma}^T \mathbf{z}_i)} = \sigma_{0u} \cdot \text{IVM}_{0i} \,. \tag{8}$$

A final methodological note concerns the panel structure of our data. The fact that respondents report up to five WTP bids for five out of 14 different electricity mixes affords an opportunity to estimate the associated parameters using fixed-effects panel estimation methods. Employing the common least-squares dummy variable (LSDV) approach, however, would not serve to identify the coefficients of the respondent attributes, as these do not vary over responses and hence would drop out of the estimation. We consequently apply an estimation method suggested by FRONDEL and VANCE (2010) that is based on the within-group estimator. It involves producing fixed-effects estimates by way of demeaning the respondent-varying variables $\mathbf{mix}_{it}$ and using Ordinary Least Squares (OLS):

$$WTP_{it} = \alpha_0 + (\boldsymbol{\alpha}_{\mathbf{mix}})^T \cdot [\mathbf{mix}_{it} - \overline{\mathbf{mix}}_i] + \boldsymbol{\alpha}^T \mathbf{z}_i + \xi_i + \nu_{it}, \tag{9}$$

where $\mathbf{mix}_{it}$ denotes a vector of 14 dummy variables indicating the concrete power mixes shown to respondent i at time t, $\boldsymbol{\alpha}_{\mathbf{mix}}$ are the related parameters, and bars denote means over the five responses of respondent i. In contrast to the LSDV approach, both sets of regressors, $\boldsymbol{\Delta}_{it}$ and $\mathbf{z_i}$, can be included simultaneously in specification (9), an advantage that becomes relevant when interaction terms involving variables from both sets are employed.

4 Empirical Results

Exploiting the panel structure of our data base, we first estimate a fixed-effects model that combines all observations, irrespective of the certainty outcome, thereby providing for a comparison with the results obtained from the switching regression model (Table 5). Not surprisingly, the signs on all of the dummies for the various power mixes are negative, indicating that the base option of 100% renewable electricity commands the highest WTP. Consistent with the descriptive statistics reported in Table 3, the strongest negative effect is seen for 100% nuclear power. With a coefficient of -63.2, the WTP for this option is about €63 lower than for electricity produced from 100% renewables. For the remaining options that include nuclear power as part of the mix, average WTPs are also lower than for those mixes that exclude nuclear power.

Turning next to the estimates of the switching regression model and beginning with the first-stage probit estimates reported in Table B1 of Appendix B, several variables have plausible and statistically significant effects on an individual's certainty about WTP bids. Most notably, the identifying variable indicating knowledge of the broad range of the levy for renewables, denoted by L, has a positive impact, as do membership in either of the treatment groups indicated by *cheap talk* or *consequential*. Conversely, female has a negative association. Likewise, relative to the base category of low-income households, the coefficients of all the income categories are uniformly negative, indicating that more wealthy households are less likely to be sure about their WTP responses. Finally, age has a non-linear effect that takes the form of an inverted

U-curve, peaking at an age of about 49.

With regard to the second-stage switching regression outcomes on WTP presented in the final columns of Table 5, several outcomes bear highlighting. First, the WTP for many electricity mixes is lower among respondents who are definitely certain about their bid (Regime 1: $C_i = 1$), with statistically significant differences obtained for several electricity mixes, such as that including 25% fossil fuels and 75% renewables. Likewise, comparing the constant terms across the two regimes reveals that the WTP for 100% renewables among definitely certain respondents who have not received any treatment is about €49.2 lower than those not reporting definite certainty (Regime 0: $C_i = 0$), a difference that is statistically significant. In contrast, as in the case of the comparison model, both the interaction terms of consequential treatment and cheap talk with the case of 100% renewables are not statistically significant, indicating that these treatments do not reduce the WTP for electricity that is exclusively produced on the basis of renewable technologies.

Three further discrepancies become evident between the two certainty regimes: First, the coefficient of the female dummy is statistically significant only for Regime 0 and over three times the magnitude of the comparison model, indicating substantially lower WTP among definitely certain females. Second, of the remaining personal attributes, the dummies indicating the income categories are positive throughout for Regime 0 and statistically significant, whereas for Regime 1 only the dummy variable corresponding to households with very high incomes is statistically positive, suggesting that wealthy households are prepared to pay more for electricity irrespective of its kind of production. Third, the coefficient on the inverse Mills ratio is negative and statistically significant only for sub-sample $C_i = 0$, suggesting that unobservables that increase the likelihood of membership in this regime have a negative effect on WTP.

Table 5: Fixed-Effects and Switching Regression Results.

	Fixed Effects		2. Stage of Switching Regression				Tests on
	Total Sample		Sub-sample $C_i = 1$		Sub-sample $C_i = 0$		Differences
		Std.		Std.		Std.	χ^2
Variable	Coeff.s	Errors	Coeff.s	Errors	Coeff.s	Errors	Statistics
Δ *75%fos25%ren*	**-20.4	(3.01)	**-25.2	(5.01)	**-16.0	(3.13)	2.43
Δ *50%fos50%ren*	**-11.8	(2.23)	**-15.5	(2.89)	* -8.52	(3.31)	2.54
Δ *25%fos75%ren*	* -13.0	(5.89)	**-25.3	(3.80)	0.58	(12.0)	* 4.26
Δ *75%fos25%nuc*	**-35.6	(4.14)	**-36.1	(7.16)	**-35.9	(3.97)	0.00
Δ *50%fos50%nuc*	**-41.6	(2.65)	**-46.9	(2.57)	**-36.4	(4.73)	* 3.79
Δ *25%fos75%nuc*	**-50.4	(2.33)	**-52.9	(2.98)	**-48.3	(3.67)	0.96
Δ *100%nuc*	**-63.2	(3.17)	**-65.3	(3.98)	**-61.8	(5.10)	0.30
Δ *25%ren75%nuc*	**-35.5	(5.95)	**-33.7	(10.5)	**-38.3	(4.66)	0.16
Δ *50%ren50%nuc*	**-31.3	(2.75)	**-34.7	(4.23)	**-28.3	(3.56)	1.36
Δ *75%ren25%nuc*	**-18.2	(2.13)	**-24.5	(2.44)	**-11.8	(3.56)	** 8.67
Δ *50%fos25%ren25%nuc*	**-23.5	(4.86)	**-31.7	(2.42)	-14.9	(9.91)	2.73
Δ *25%fos50%ren25%nuc*	**-20.2	(3.23)	**-21.7	(4.92)	**-19.3	(4.06)	0.14
Δ *25%fos25%ren50%nuc*	**-36.5	(2.53)	**-40.4	(2.58)	**-33.1	(4.54)	1.95
female	**-6.82	(1.87)	-4.63	(3.25)	**-22.8	(7.94)	* 4.51
age	-0.11	(0.08)	-0.07	(0.05)	-0.16	(0.14)	0.49
cheap talk	4.19	(2.73)	4.03	(2.43)	13.5	(7.79)	1.36
cheap talk ×Δ *100%ren*	-4.61	(4.69)	-4.53	(3.17)	-8.92	(10.0)	0.37
consequential	* 5.93	(2.71)	* 7.78	(3.20)	8.96	(5.02)	0.05
consequential ×Δ *100%ren*	0.91	(2.32)	-4.33	(3.12)	1.94	(3.34)	1.88
degree	0.80	(2.00)	2.01	(2.02)	1.32	(3.67)	0.03
children	-2.61	(4.62)	-3.80	(2.19)	1.89	(10.9)	0.26
medium income	* 5.62	(2.20)	3.54	(3.18)	** 9.60	(3.33)	1.75
high income	** 11.6	(3.73)	9.33	(5.14)	** 20.2	(5.8)	1.95
very high income	** 19.9	(7.56)	* 14.6	(5.79)	* 43.6	(20.6)	1.85
IVM_1	–	–	-0.62	(15.5)	–	–	–
IVM_2	–	–	–	–	*-54.0	(21.8)	–
const.	**91.9	(4.42)	** 89.4	(12.5)	**138.6	(18.0)	* 5.03
Number of Obs.	24,906		13,310		11,596		

To glean further insight into the implications of endogenous switching for the parameter estimates, we also estimated two models that exclude the inverse Mills ratio (see Table B2 in Appendix B). Overall, the differences in the coefficient estimates between the two regimes is less stark when not controlling for sample selectivity. Most notably, the WTP estimates for the case of 100% renewable energy, captured by the constant terms, are statistically indistinguishable, indicating that the application of the switching regression model identifies differences between definitely and probably certain respondents that would otherwise be obscured when not controlling for the endogeneity of certainty status.

5 Summary and Conclusion

Various techniques, such as consequential- and cheap-talk scripts, as well as the certainty approach, have been proposed in the literature to mitigate the well-known bias in hypothetical responses to questions on the willingness-to-pay (WTP) for non-market goods such as environmental amenities. Using an experimental design and empirical data from a recent survey among about 6,500 German households on the WTP for a variety of 14 kinds of electricity mixes, we have provided further evidence on the effectiveness of these three approaches to calibrating hypothetical WTP responses. Employing an endogenous switching regression model to account for the endogeneity of respondent certainty, we have applied the certainty approach on continuous WTP bids, whereas it has been used only for dichotomous choice questions by, e. g. , BLUMENSCHEIN et al. (2008).

Our results suggest that neither cheap talk nor the consequential script fundamentally bear on the estimates of WTP. However, these treatments are found to have an indirect effect by virtue of their positive influence in increasing the probability that the respondent is certain of their estimate. By contrast, when controlling for selectivity bias with a switching regression model, we uncover differences in WTP according to respondent certainty, with the certain respondents reporting lower values for several kinds of electricity mixes, most notably the 100% renewable mix. This raises the question of how to treat the WTP estimates derived from the two groups.

BLUMENSCHEIN et al. (2008) discard the values obtained from the uncertain respondents, persuasively arguing that an unsure response is tantamount to a no-answer to a dichotomous choice question, as those uncertain about a purchase are unlikely to actually spend money even when they state a tentative willingness to do so. When following this approach also in cases in which respondents indicate their WTP on a continuous scale, as in the present study, policy recommendations that take heed of the, on average, lower WTP of those who are definitely certain about their responses would seem warranted. To this end, the switching regression provides a useful methodology for controlling for selectivity biases as related to the certainty of the respondents.

Appendix A: Extract of Questionnaire

The elicitation of the WTP for specific electricity mixes began with a brief introduction on the diversity of production technologies, followed by a short description of the survey design, including several practical examples. Upon displaying the introductory text, both the cheap-talk and consequential scripts were presented to the respective treatment groups before posing the question on WTP, yet not to the control group. The translations of these texts and scripts into English is reported below:

Introductory text

(READ BY ALL PARTICIPANTS)

Electricity can be produced with different energy sources and technologies. Among these are coal- or natural gas fired power plants, nuclear power, or renewable energy technologies such as photovoltaics, hydropower, and wind turbines. A household might obtain electricity that is produced from a single source such as a fossil fuel, or it might alternatively obtain electricity that is produced from some mix of different sources such as fossil fuels, nuclear power, and renewable energies.

We will now present you with different electricity offers that are distinguished solely by the proportions of fossil fuels, nuclear energy, and renewable energy with which the electricity is produced. For each of these offers, we request that you report the maximum amount that you, personally, would be willing to pay. As a basis for comparison, please consider an energy mix comprised exclusively of the fossil sources coal, natural gas, and oil, which has a price of €100 per month.

Example: The price for the comparison offer is €100. If the price you would be willing to pay for the alternative offer were €70, please record the amount €70. If the price you would be willing to pay for the alternative offer were instead €180, please record the amount €180. Of course, any other values may also be recorded.

Willingness to pay for alternative energy sources

(READ BY ALL PARTICIPANTS)

Now we would like to ask you about how much you would be willing to pay for different energy sources and energy technologies. In what follows, we will refer to this as your "willingness to pay."

CHEAP TALK (READ BY GROUP 2):

In analyzing survey data it is often found that some respondents report a relatively high willingness to pay for environmental goods like clean air. Presumably, these respondents don't take into account that were they really to pay such a large sum of money, they would have to forgo the purchase of other goods. We therefore request that your answer to the following questions corresponds to the sum of money that you would in reality be willing to pay.

CONSEQUENTIAL SCRIPT (READ BY GROUP 3):

First we would like to ask whether you believe that representative surveys such as this one have an influence on policy-making. More concretely: do you believe that the results of this survey will have an influence on political decisions?

- Yes
- No
- Don't know

Irrespective of your answer to the above question, we would like you to assume that this representative survey will have an influence on political decisions. This means that your reported willingness to pay should correspond to the amount of money that you are truly willing to spend.

Appendix B: Tables

Table B1: First-Stage Probit Estimation Results.

	Probit Estimation		Marginal Effects	
Variable	Coeff.s	Std. Errors	Effects	Std. Errors
female	** -0.350	(0.039)	** -0.135	(0.015)
age	** 0.056	(0.009)	** -0.002	(0.001)
age × *age*	**$-5.6 \cdot 10^{-4}$	$(9.0 \cdot 10^{-5})$	–	–
cheap talk	** 0.235	(0.043)	** 0.090	(0.016)
consequential	** 0.126	(0.043)	** 0.048	(0.016)
degree	* 0.092	(0.039)	* 0.035	(0.015)
children	0.029	(0.066)	0.011	(0.025)
medium income	0.022	(0.052)	0.008	(0.020)
high income	0.091	(0.062)	0.035	(0.024)
very high income	** 0.291	(0.077)	** 0.111	(0.029)
2-person hh	* -0.105	(0.047)	* -0.040	(0.018)
3-person hh	-0.045	(0.069)	-0.017	(0.026)
4-person hh	-0.118	(0.086)	-0.045	(0.033)
> *4-person hh*	-0.043	(0.115)	-0.017	(0.044)
P	0.083	(0.046)	0.032	(0.017)
L	** 0.229	(0.039)	** 0.088	(0.015)
const.	** -1.364	(0.230)	–	–
Number of Obs.		5,283		

Table B2: Fixed-Effects Results without Switching Regression Correction.

	Sub-sample $C_i = 1$		Sub-sample $C_i = 0$		Test on Differences
		Std.		Std.	χ^2
Variable	Coeff.s	Errors	Coeff.s	Errors	Statistics
Δ *75%fos25%ren*	**-25.2	(5.01)	**-15.7	(3.13)	2.57
Δ *50%fos50%ren*	**-15.5	(2.89)	* -8.22	(3.33)	2.72
Δ *25%fos75%ren*	**-25.4	(3.80)	0.83	(12.0)	* 4.33
Δ *75%fos25%nuc*	**-36.1	(7.15)	**-35.6	(3.92)	0.00
Δ *50%fos50%nuc*	**-46.9	(2.57)	**-36.2	(4.77)	* 3.93
Δ *25%fos75%nuc*	**-52.9	(2.98)	**-48.1	(3.68)	1.07
Δ *100%nuc*	**-65.3	(3.99)	**-61.6	(5.09)	0.34
Δ *25%ren75%nuc*	**-33.7	(10.5)	**-38.0	(4.74)	0.14
Δ *50%ren50%nuc*	**-34.7	(4.23)	**-27.9	(3.54)	1.50
Δ *75%ren25%nuc*	**-24.5	(2.44)	**-11.6	(3.59)	8.96
Δ *50%fos25%ren25%nuc*	**-31.7	(2.42)	-14.6	(9.97)	2.78
Δ *25%fos50%ren25%nuc*	**-21.7	(4.92)	**-19.1	(4.07)	0.17
Δ *25%fos25%ren50%nuc*	**-40.4	(2.58)	**-33.1	(4.57)	2.10
female	*4.76	(2.41)	**-10.4	(3.42)	1.83
age	-0.09	(0.05)	-0.17	(0.14)	0.28
cheap talk	4.07	(2.62)	5.48	(5.51)	0.05
cheap talk $\times\Delta$ *100%ren*	-2.53	(3.16)	-8.47	(10.0)	0.32
consequential	* 7.71	(3.88)	4.49	(3.70)	0.36
consequential $\times\Delta$ *100%ren*	-4.34	(3.12)	2.22	(3.36)	2.05
degree	1.99	(2.41)	0.33	(3.41)	0.31
children	-3.77	(2.12)	0.76	(10.3)	0.09
medium income	3.53	(3.21)	* 7.71	(3.01)	0.90
high income	9.41	(5.59)	** 14.6	(4.99)	0.47
very high income	14.7	(8.11)	29.8	(15.9)	0.71
const.	** 88.9	(4.00)	** 95.6	(7.6)	0.61
Number of Obs.	13,310		11,596		

References

BATEMAN, I. J. , LANGFORD, I. H. (1997) Budget constraint, temporal and ordering effects in contingent valuation studies. *Environment and Planning A* 29, 1215-1228.

BLUMENSCHEIN, K. , BLOMQUIST, G. C. , JOHANNESSON, M. , HORN, N. , FREEMAN, P. R. (2008) Eliciting Willingness to Pay in the Without Bias: Evidence from a Field Experiment. *The Economic Journal* 118(525), 114-137.

BLUMENSCHEIN, K. , JOHANNESSON, M. , BLOMQUIST, G. C. , LILJAS, B. , O'CONNOR, R. M. (1998) Experimental Results on Expressed Certainty an Hypothetical Bias in Contingent Valuation. *Southern Economic Journal* 65(1), 169-177.

BLUMENSCHEIN, K. , JOHANNESSON, M. , YOKOYAMA, K. K. , FREEMAN, P. R. (2001) Hypothetical versus Real Willingness to Pay in the Health Care Sector: Results from a Field Experiment. *Journal of Health Economics* 20(3), 441-457.

BULTE, E. , GERKING, S. , LIST, L. A. , DE ZEEUW, A. (2005) The Effect of Varying the Causes of Environmental Problems on Stated WTP Values: Evidence from a Field Study. *Journal of Environmental Economics and Management* 49(2), 330-342.

CARLSSON, F. , FRYKBLOM, P., LAGERKVIST, C. J. (2005) Using Cheap Talk as a Test of Validity in Choice Experiments. *Economics Letters* 89(2), 147-152.

CARSON, R. , GROVES, T., LIST, J. A. , MACHIN, M. (2004) Probabilistic Influence and Supplemental Benefits. A Field Test of the Two Key Assumptions Underlying Stated Preferences. Unpublished manuscript.

CLARK, J. , FRIESEN, L. (2008) The Causes of Order Effects in Contingent Valuation Surveys: an Experimental Investigation. *Journal of Environmental Economics and Management* 56(2), 195-206.

CUMMINGS, R. G. , ELLIOT, S. , HARRISON, G. W. , MURPHY, J. (1997) Are Hypothetical Referenda Incentive-Compatible? *Journal of Political Economy* 105(3), 609-621.

CUMMINGS, R. G. , HARRISON, G. W. , RUTSTRÖM, E. E. (1995) Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive-Compatible? *American Economic Review* 85(1), 260-266.

CUMMINGS, R. G. , TAYLOR, L. O. (1999) Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method. *American Economic Review* 89(3), 649-665.

FREW, E. J. , WHYNES, D. K. , WOLSTENHOLME, J. L. (2003) Eliciting Willingness to Pay: Comparing Closed-Ended with Open-Ended and Payment Scale Formats. *Medical Decision Making* 23(2), 150-159.

FRONDEL, M., VANCE, C. (2010) Fixed, Random, or Something in Between? A Variant of Hausman's Specification Test for Panel Data Estimators. *Economics Letters* 107, 327-329.

GRÖSCHE, P. , SCHRÖDER, C. (2011) Eliciting Public Support for Greening the Electricity Mix Using Random Parameter Techniques. *Energy Economics* 33 (2), 363-370.

HALSTEAD, J. , LULOFF, A. , STEVENS, T. H. (1992) Protest Bidders in Contingent Valuation. *Northeastern Journal of Agricultural and Resource Economics* 21 (2), 160-169.

HARRISON, G. W. (2006) Experimental Evidence on Alternative Environmental Valuation Methods. *Environmental and Resource Economics* 34 (1), 125-162.

HARRISON, G. W. , HARSTAD, R. , RUTSTRÖM, E. E. (2004) Experimental methods and elicitation of values. *Experimental Economics* 7(2): 123-140.

HARRISON, G. W. , RUTSTRÖM, E. E. (2008) Experimental evidence on the existence of hypothetical bias in value elicitation methods. In: C. Plott and V. L. Smith (eds.), *Handbook of Experimental Economics Results*, Edition 1, Volume 1, Number 5. New York: Elsevier Science.

JOHANNESSON, M. , LILJAS, B. , JOHANSSON, P-O. (1998) An Experimental Comparison of Dichotomous Choice Contingent Valuation Questions and Real Purchase Decisions. *Applied Economics* 30(5), 643-647.

LANCASTER, K. J. , (1966) A New Approach to Consumer Theory. *Journal of Political Economy* 74, 132-157.

LIST, J. A. (2001) Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sportscards. *American Economic Review* 91(5), 1498-1507.

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

MADDALA, G. S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*. Reprint 1999. Cambridge University Press, Cambridge.

MENGES, R. , SCHRÖDER, C. , TRAUB, S. (2005) Altruism, Warm Glow and the Willingness-to-Donate for Green Electricity: an Artefactual Field Experiment. *Environmental and Resource Economics* 31(4), 431-458 .

MURPHY, J. J. , STEVENS, T. , WEATHERHEAD, D. (2005) Is cheap talk effective at eliminating hypothetical bias in a provision point mechanism? *Environmental and Resource Economics* 30(3), 327-343.

WHITEHEAD, J. C. , CHERRY, T. L. (2007) Willingness to Pay for a Green Energy Program: A Comparison of Ex-ante and Ex-post Hypothetical Bias Mitigation Approaches. *Resource and Economics* 29(4), 247-261.