WORKING PAPER SERIES

Remanufacturing and Consumers' Risky Choices: Behavioral Modeling and the Role of Ambiguity Aversion

James D. Abbey/Rainer Kleber/Gilvan C. Souza/Guido Voigt

Working Paper No. 11/2018

OTTO VON GUERICKE
UNIVERSITÄT
MAGDEBURG

FACULTY OF ECONOMICS
AND MANAGEMENT

Impressum (§ 5 TMG)
Herausgeber:
Otto-von-Guericke-Universität Magdeburg
Fakultät für Wirtschaftswissenschaft
Der Dekan

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

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

Bezug über den Herausgeber
ISSN 1615-4274

Remanufacturing and Consumers' Risky Choices: Behavioral Modeling and the Role of Ambiguity Aversion[1]

James D. Abbey
Mays Business School,
Texas A&M University, College Station, TX, jabbey@mays.tamu.edu

Rainer Kleber
Faculty of Economics and Management,
Otto-von-Guericke University, Magdeburg, Germany, rainer.kleber@ovgu.de

Gilvan C. Souza
Kelley School of Business,
Indiana University, Bloomington, IN 47405, gsouza@indiana.edu

Guido Voigt
Logistics and Supply Chain Management,
University of Hamburg, Hamburg, Germany, guido.voigt@uni-hamburg.de

Abstract

Willingness to pay (WTP) is known to be lower for remanufactured products than for comparable new products. Normative work to date has assumed that a consumer's WTP for a remanufactured product is a fraction, called discount factor, of the consumer's WTP for a corresponding new product, and that this discount factor is constant across consumers. Recent empirical research demonstrates, however, that the discount factor is not constant across consumers. This discovery has led researchers to call for an exploration of more refined utility models that incorporate heterogeneous risk preferences through elements such as risk aversion, loss aversion, and ambiguity aversion. To address this call, this manuscript assesses each of these risk preference elements by empirically deriving WTP distributions from two interlinked studies. To provide triangulation in both the empirical method and sample, the interlinked studies employ an online survey and a laboratory experiment that elicits WTP for framed lotteries that proxy the situation of buying remanufactured products. The empirical results and robustness verifications demonstrate that a parsimonious standard utility model incorporating only risk aversion explains the WTP data reasonably well.

Keywords: Closed-loop supply chains, decision heuristics and decision rules, remanufacturing, behavioral operations

1 Introduction

Remanufacturing is the process of restoring a used product to a common aesthetic and operating standard, which is used in many industries (Lund, 1984; ANSI Consensus Body, 2015). In many

[1] The paper has been accepted for publication in the **Journal of Operations Management**.

cases, the products remanufactured by the original equipment manufacturer (OEM) are, from a technical perspective, effectively like new (Ferguson et al., 2006). However, there is significant empirical evidence that consumers are not willing to pay the same amount for remanufactured products as for new products, which is reflected in observed prices. For instance, Subramanian and Subramanyam (2012) compare prices of new and remanufactured products at eBay, and find that price discounts for remanufactured products, after controlling for warranty, range between 15% and 40%, depending on the product category. Guide and Li (2010) auctioned new and remanufactured power tools and Internet routers on eBay, with the same warranty and with fulfillment guaranteed by the respective OEMs. They find that the final auction price for remanufactured products is between 10% and 15% lower than for their corresponding new counterparts. To model consumer behavior towards remanufactured and new products, consider a consumer base that is vertically differentiated in their willingness to pay (WTP) for a new product (e.g., a particular model of a power tool). The closed-loop supply chain (CLSC) literature models a consumer's WTP for the corresponding remanufactured product as a fraction δ of his/her WTP for the new product, where $\delta < 1$ is assumed to be constant across consumers for parsimony. A considerable amount of normative research has been conducted in the CLSC literature using this assumption (for a review, see Souza, 2013).

Recent research, however, empirically demonstrates that the discount factor δ is not constant across consumers (Abbey et al., 2017). Essentially, consumers view buying a remanufactured product as a risky decision, with considerable ambiguity regarding the quality of the remanufactured product (Ovchinnikov, 2011; Abbey et al., 2015b, 2017). In particular, Abbey et al. (2017) empirically find that a consumer's own δ can be explained by his/her *perception* of quality, of which *perceived* probability (or risk) of cosmetic and functional defects are two dimensions. The behavioral literature outlines three primary dimensions of risk preferences, whose magnitude differ among consumers and may also be context dependent: risk aversion, ambiguity aversion, and loss aversion. These concepts are discussed in more depth in section 2. In the context of this study, risk refers to a consumer's *own* assessment of the likelihood of defects in a remanufactured product, knowing that the input for the remanufacturing process (i.e., used products) is variable. Ambiguity refers to the inability of consumers to resolve the exact meaning of remanufacturing—how thorough is the remanufacturing process? Loss aversion refers to weighing potentially poor quality outcomes (i.e., defects) much stronger than the prospect of receiving a product that is indeed as good as new. We note that risk, loss, and ambiguity aversion differ across consumers, and for a specific consumer they may also be different across product types, such as mobile phones or household goods.

The goal of this research is to propose a theoretically grounded, empirically validated, and reasonably parsimonious model refinement of WTP for remanufactured products. This parsimonious WTP model accounts for risk aversion in conjunction with the CLSC literature's findings that consumers view the purchase of remanufactured products as a risky decision. To this end, we conduct a combination of two empirical studies, using both an online survey and controlled laboratory experiments, to collect data on WTP for remanufactured mobile phones. In the lab experiments, there is a mechanism to control for WTP for new phones, so that the resulting variability in WTP is mostly explained by the variability in δ, in order to build on the results of Abbey et al. (2017). Our decision to use a technology product for the study follows the finding by Abbey et al. (2015b, 2017) that this product category was less likely than others to engender aversions, such as feelings of disgust among consumers. We then fit standard utility models from the behavioral literature to our WTP data: a complete model with risk, ambiguity, and loss aversion; a subset of this model with risk and loss aversion only; finally a subset with risk aversion only. Interestingly, we find that the utility model that considers only risk aversion explains the empirically observed WTP distribution sufficiently well. Adding loss and ambiguity aversion to the model does not significantly improve the fit. In the coming section, we elaborate on our contribution within the domains of risky-choice and CLSC literature.

2 Literature Review

This paper is related to two streams of literature: CLSC research and risky-choice theory—decision making under risk and ambiguity. We review each stream separately.

2.1 CLSC Literature

Despite the lower consumer WTP for remanufactured products, OEMs remanufacture consumer returns for reasons such as value recovery, product line extension, brand protection against the possibility of a low-quality third-party firm remanufacturing the OEM's product, and cannibalization of competitors' product sales (see, e.g., Atasu et al., 2010; Ferguson and Souza, 2010). Such reasons led to a considerable stream of prescriptive CLSC research on an OEM's decision to remanufacture, including pricing and competition (e.g., Debo et al., 2005; Ferguson and Toktay, 2006; Ferrer and Swaminathan, 2006), marketing strategies (e.g., Atasu et al., 2008; Oraiopoulos et al., 2012), impact of remanufacturing on new product design (Akturk et al., 2017; Atasu and Souza, 2013; Galbreth et al., 2013; Subramanian et al., 2013; Örsdemir et al., 2014), and production capacity (Vorasayan and Ryan, 2006). In this research, the WTP discount factor for remanufactured products is assumed to be constant across consumers, as previously discussed.

Researchers studying behavioral aspects of closed-loop supply chains are now extending their investigation of the discount factors. Ovchinnikov (2011) finds that consumers perceive remanufactured laptops as inferior in quality if the associated discount levels are too high. Using lab experiments, Agrawal et al. (2015) find that the presence of remanufactured Apple iPod MP3 players may cause consumers to lower their WTP for the respective new MP3 players, as consumers infer a high level of returns due to low quality. Ovchinnikov et al. (2014) empirically derive demand functions for new and remanufactured products in the context of a firm that offers a product-service bundle, such as AT&T for cellular service. Remanufactured technology products were shown to increase in attractiveness at a decreasing rate as the price discounts increased, but—in line with Ovchinnikov (2011)—excessive price discounts could negatively impact attractiveness (Abbey et al., 2015b). Abbey et al. (2017) demonstrate that perceived risk of functional defects and cosmetic defects are two main factors underlying lower WTP for remanufactured electronic products. Since quality risk perceptions are inherently personal traits, they hypothesize and empirically confirm that consumers differ significantly in their discount factors for remanufactured products. Table 1 summarizes the recent behavioral CLSC literature and the fit of this manuscript within the evolving domain of behavioral CLSCs. Taken as a whole, an implicit theme in this research is that consumers perceive the decision to purchase a remanufactured product as a risky choice, driven by idiosyncratic risk preferences. We build upon these insights by deriving WTP distributions for framed lotteries that capture the essential trade-offs of buying a remanufactured product. We then fit our data to standard utility models from risky-choice theory, which we review next.

2.2 Risk preferences and loss aversion

The extent to which decision makers are willing to take risks (with known probabilities) constitute their risk preferences. The leading model to explain behavior under risk is that decision makers maximize expected utility (Von Neumann and Morgenstern, 1947). If the utility function is concave (convex), subjects show risk averse (seeking) behavior, that is, their certainty equivalent of a gamble is lower (higher) than the expected value of the lottery. Significant research effort has been conducted to elicit individual risk preferences (see Charness et al., 2013, for a comprehensive comparison of different elicitation methods). While the standard form of expected utility maximization with risk aversion is widely applied in economic contexts, it is often found that it does not always organize empirical data sufficiently well (see Starmer, 2000, for a review of the empirical findings).

Table 1: Most Relevant Behavioral CLSC Literature: Key Insights and Theoretical Grounding

Reference	Method	Focal Consumer Product(s)	Key Insights
Guide and Li (2010)	Field Experiment	Skil Jigsaw	Significantly lower WTP for remanufactured (vs. new) product sold on eBay. Distinct consumer segment preferences for new vs. remanufactured products.
Ovchinnikov (2011)	Lab Experiment	Dell Laptop	Significantly lower WTP for remanufactured product. Inverted-U shape relationship between demand and price discounts due to price-quality inference.
Subramanian and Subramanyam (2012)	Field Study	Consumer Electronics on eBay	Significantly lower WTP for remanufactured consumer electronics product. Seller reputation impacts WTP and consumer behavior.
Ovchinnikov et al. (2014)	Lab Experiment	Apple iPhone	Significantly lower WTP for remanufactured product. Cannibalization of new product sales.
Abbey et al. (2015b)	Lab Studies and Online Field Experiments	Various Product Categories	Multiple behavioral drivers identified. Dominant impact of perceived quality. Impact of other perceptual drivers (such as disgust) minimized in technology (electronics) category.
Abbey et al. (2015a)	Online Field Experiment	Various Product Categories	Isolation of distinct consumer segments and averse consumers (to reman. products). Segment impacts minimized in technology product category.
Agrawal et al. (2015)	Lab Experiment	Apple iPod	Significantly lower WTP for remanufactured product. Significant impact on WTP due to presence of OEM versus non-OEM remanufacturing source.
Abbey et al. (2017)	Online Studies and Lab Experiment	Apple iPod, iPhone, MacBook	Quantified and isolated impact of heterogeneity in WTP. Perceived quality as dominant WTP driver. Impact of other behavioral covariates insignificant for focal Apple products.

One factor that explains deviations from (standard) expected utility is loss aversion. Loss aversion acknowledges that decision makers have a reference point to which they ex-ante compare potential gains and losses in risky situations. With loss aversion, losses are found to have a relatively higher weight than gains. In other words, the utility function is steeper for losses than for gains and has a kink around the reference point (Köbberling and Wakker, 2005). There is a large body of literature that finds evidence for loss aversion (for influential studies, see Kahneman et al., 1990; Tversky and Kahneman, 1991).

One critical parameter to be calibrated under loss aversion is the reference point. While it seems natural to take zero payoffs as a reference point, there is substantial evidence that other reference points that are, for example, determined by past experience (e.g., a mixture of worst-case and best-case outcomes) might serve as the reference for perceiving gains and losses (Kőszegi and Rabin, 2006). Such reference points can, for example, explain over- and under-ordering behavior in newsvendor experiments (Long and Nasiry, 2014).

2.3 Ambiguity preferences

Consider an urn that contains black and red balls, and one can bet on a ball color and win a prize if the color is drawn. According to Ellsberg (1961), the decision is called to be a risky decision if the number of colored balls is known (e.g., 50 black and 50 red) and an ambiguous decision if the number of colored balls is unknown (e.g., the urn contains 100 balls but the exact numbers of black and red balls are not known).

One interesting aspect of decisions involving ambiguous probabilities is the Ellsberg-Paradox (Ellsberg, 1961), which shows that the axioms of (subjective) expected utility theory are violated when the odds of winning are ambiguous. Several models that generalize (subjective) expected utility have been proposed, among others: Choquet expected utility (Schmeidler, 1983), maxmin expected utility (Gilboa and Schmeidler, 1989), and α-MEU (Ghirardato et al., 2004). These models nest the case of known probabilities (risk) and are therefore more potent in explaining observed behavior. Yet, the increased descriptive validity comes with a higher complexity (more parameters). We refer to Etner et al. (2012) for a thorough discussion of recent theoretical and empirical developments, and to Camerer and Weber (1992) and Camerer (1995) for comprehensive and insightful reviews on ambiguity.

Ellsberg's experiments also laid the foundation for considering ambiguity preferences as a part of utility functions, since subjects in the original Ellsberg experiments preferred to take risk in situations where they knew the exact probabilities rather than in situations where these probabilities were unknown.

In a series of studies, researchers showed that depending on the context (e.g., urn type) and the odds of winning, individual's ambiguity preferences range from being ambiguity averse, neutral or even seeking (Abdellaoui et al., 2011; Einhorn and Hogarth, 1988; Cohen et al., 1987; Curley and Yates, 1989; Fox and Tversky, 1998; Ghosh and Ray, 1997; Halevy, 2007; Smith et al., 2002; Trautmann et al., 2011). From Fox and Weber (2002) and Heath and Tversky (1991), it is well known that different decision frames can considerably affect the decision makers' ambiguity preferences. Gneezy et al. (2015) show that the degree of ambiguity aversion is likely to be overstated if risk neutrality is assumed (i.e., if ambiguity preferences and risk preferences are not elicited and estimated jointly).

There is evidence that subjects are more likely to exhibit ambiguity aversion when comparing lotteries with differing degrees of ambiguity instead of evaluating each lottery separately (see Chow and Sarin, 2001; Fox and Tversky, 1995; Fox and Weber, 2002; Trautmann and Schmidt, 2012). This observation is known as the comparative ignorance effect. In our context, this means that subjects' WTP is only strongly affected by ambiguity aversion if subjects directly compare remanufactured devices with different degrees of ambiguity. For example, this would occur if information on the quality of remanufactured devices is available in online forums for one brand but not for another. In this example, ambiguity aversion would not strongly affect subjects' WTP if they are comparing different models of the same brand.

Within this vast body of literature on ambiguity preferences, our study is more related to research that applies and empirically tests theoretical constructs to a specific economic context (e.g., Ahn et al., 2014; Chen et al., 2007; Ho et al. 2002; Muthukrishnan et al. 2009). In our case, the theoretical constructs are utility models incorporating risk, ambiguity, and loss aversion, while the context is that of consumer behavior towards remanufactured products.

Camerer and Weber (1992) argue that, from a purely theoretical perspective, outcome ambiguity is too coarse a category because it can be interpreted as an unknown probability distribution over a potential range of outcomes. This, in turn, translates to probabilistic ambiguity. Thus, in our study we model ambiguity through probabilistic ambiguity. Specifically, we have treatments where we announce an exact 25% probability of buying a remanufactured phone with defects (only risk, no ambiguity), and treatments where we announce that the probability of buying a defective phone is between 0 and 50% (ambiguity). We contribute to this stream of literature by conducting controlled laboratory experiments in which we vary exogenously the degree of ambiguity in probabilities in a decision involving remanufactured products.

2.4 Discussion

Significant research progress has been made to disentangle the different facets of decisions under risk and uncertainty. There is a large body of research that aims at modeling, eliciting, and estimating preferences concerning risk, losses, and ambiguity. The results indicate that preferences are dependent on the stake size, the odds of winning, and the economic context. However, the downside of more complex behavioral models is that there are several parameters that need to be estimated in the decision context under investigation. This requires large samples due to many degrees of freedom (hence, over fitting is a concern), and subtle elicitation methods.

We approach the problem from a different angle that acknowledges previous research efforts to estimate risk and ambiguity preferences. We take this state-of-the-art knowledge to build a behavioral model of consumer choice for remanufactured products that explains our observed WT distribution reasonably well. By this means, we avoid over fitting our observations to the moving parts of the theory with the trade-off that other parameters and/or shapes of utility functions—some examined through various forms of robustness analysis—might also fit the data well. Nonetheless, the key observation is that a parsimonious model that only includes a (heterogeneous) risk aversion parameter, estimated by a large-scale study, has high descriptive power. This finding will hopefully encourage researchers to consider risk aversion as a key behavioral aspect to test the robustness of analytic models that assume risk neutrality and homogeneous risk preferences in issues surrounding demand for remanufactured products.

3 Behavioral experiment

Abbey et al. (2017) show that consumers have a lower WTP for remanufactured products because they perceive them to have a risk of cosmetic and/or functional defects. We use this finding in an online study that delves into the perceived probabilities of experiencing a defect and the resultant WTP for both a remanufactured Apple iPhone SE and a comparably priced Huawei P9. We employ Apple as a focal brand to maintain consistency with previous studies (e.g., Agrawal et al., 2015; Abbey et al., 2017). We also test the lesser known Huawei brand to discern potential brand effects. In a second step, we elicit in incentivized lab experiments the WTP for lotteries that capture the essential trade-offs reported in the online study. The elicited WTP distributions serve as a proxy for the WTP for the respective remanufactured mobile phones. Proxying the WTP distributions with lotteries that are tailored to real-life stakes when buying remanufactured phones is essential for our goal to control for other variables that determine WTP, with the purpose of focusing on the perceived risky nature of buying remanufactured mobile phones.

Table 2: MTurk Results: WTP for remanufactured Apple iPhone SE and Huawei P9 phones[a]

Focal product	Market price	WTP reman	Point estimate	Lower Range	Upper Range	WTP with known defect
Apple	\$399	\$197.0 (98.0)	23.9% (20.1)	11.3% (15.8)	37.3% (24.4)	\$151.7 (82.9)
Huawei	\$399	\$175.7 (89.2)	25.8% (19.6)	13.5% (16.5)	40.2% (24.2)	\$125.9 (76.4)

[a] Numbers in parenthesis are standard deviations. The data was obtained in a between-subjects design

The controlled laboratory environment also allows us to disentangle the different facets of risk preferences (risk, ambiguity, and loss aversion) and cleanly link the buying decision to risky-choice theory. Figure 1 illustrates the connection between the pre-study and the laboratory experiments.

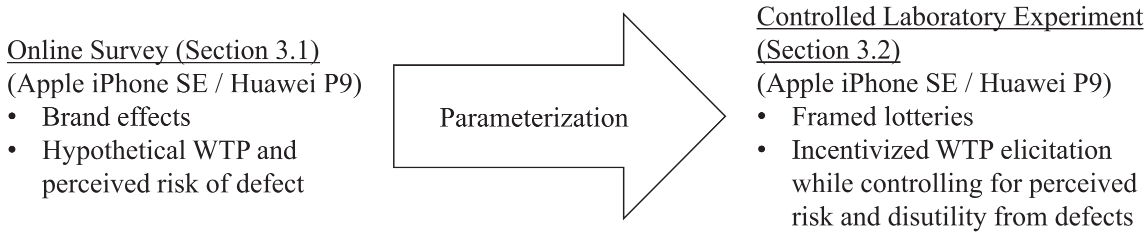


Figure 1: Outline of experiments

3.1 Pre-study on Amazon MTurk

We conducted an online survey using Amazon MTurk (`www.mturk.com`). Amazon MTurk is a commercial platform that allows a requester to outsource tasks, such as answering a survey, to participants that are paid a fee determined by the requester. Each participant was paid \$2.25 to completely answer a short survey. In the study, 123 respondents stated their WTP as well as their perceived probabilities of encountering a defect (point estimate, lower range, upper range) with two remanufactured products: an Apple iPhone SE and a Huawei P9. We selected these products because both had the same market price of \$399 when new in June 2017, and had similar technical specifications. In displaying the products to study participants, we varied how the products were presented (sequentially, in randomized order, or simultaneously) and demonstrated that how the products were presented did not affect the results ($p > 0.28$). Table 2 summarizes the results, which we elaborate on below.

Based on a market price of \$399, respondents indicated a desire for a 50.6% discount (average WTP of \$197, standard deviation (sd) of 98.0) on a remanufactured Apple iPhone SE and a 56.0% discount (average WTP of \$175.7, sd of 89.2) on a comparable remanufactured Huawei

P9. This is a highly significant within-subject difference in WTP ($p < 0.01$).

Though the WTP showed significant differences, the respondents indicated little difference in their perceived average percentage of defects across brands. Specifically, the respondents indicated that they perceived an average percentage of a defect with the remanufactured Apple product of 23.9% (sd: 20.1%) and a 25.8% (sd: 19.6) chance of a defect with the remanufactured Huawei phone; this difference is not statistically significant ($p = 0.16$).

Respondents reported an average percentage range of a defect between 11.3% (sd: 15.8) and 37.3% (sd: 24.4) for the remanufactured Apple phone, and between 13.5% (sd: 16.5) and 40.2% (sd: 24.2) for the remanufactured Huawei phone. Comparing the average lower values of these ranges (11.3% vs. 13.5%) we find a significant effect ($p < .05$). Comparing the average upper values of these ranges (37.3% vs. 40.2%) we find a weakly significant effect ($p < .10$). When subjects were told that there was a known defect (a visible scratch on the screen), the average WTP dropped significantly to \$151.7 (sd: 82.9) and \$125.9 (sd: 76.4)—drops of \$45.7 and \$49.7—for the remanufactured Apple and Huawei products, respectively. These large decreases in WTP were highly statistically significant within-subject effects ($p < 0.001$).

The results show that subjects were ambiguous regarding the perceived probabilities of receiving a defective remanufactured unit, which manifests itself in the stated probability ranges. Further, there is a substantial amount of heterogeneity in the variance of the reported probability ranges and WTP statements. In sum, the results support our view that consumers perceive buying remanufactured products as a gamble (i.e., an as-good-as-new item with a certain probability vs. a defective item with the reciprocal probability). Thus, approximating WTP distributions via framed lotteries, as seen in the laboratory experiment described below, appears to be reasonable.

3.2 Controlled laboratory experiment

A commonly used approach for eliciting WTP is the Becker-DeGroot-Marschak (BDM) method (Becker et al., 1964), which is also known to be incentive compatible. Applying the BDM method to directly elicit WTP for a remanufactured product in a laboratory experiment, one would ask a subject to first state his/her WTP, then determine a selling price from the draw of a uniform distribution with an appropriate support, and finally execute the sale of the product if the selling price turns out to be lower than the stated WTP. The approach of actually selling the product has the drawback that we cannot uniquely establish the link between consumers' WTP and consumer's risk preferences, since previous studies (including our pre-study) indicate that the WTP for the remanufactured product is, among others, a function of the WTP for the

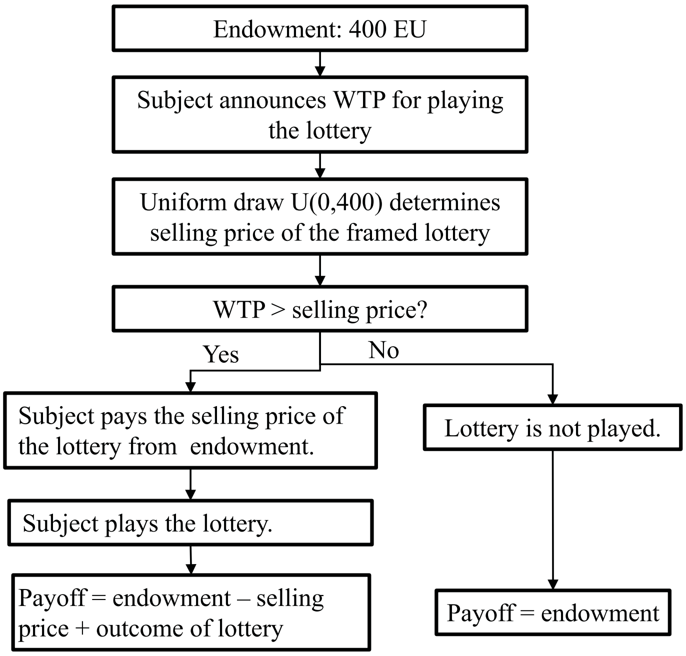
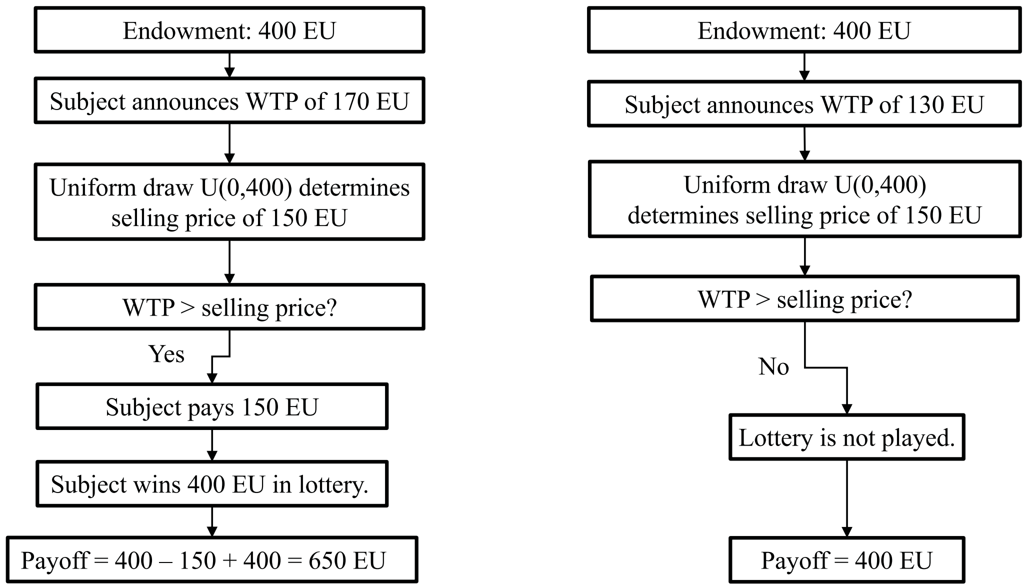


Figure 2: BDM method for framed lotteries

new product and the likelihood of/disutility from receiving a defective item. Thus, in order to cleanly isolate the effect of risk preferences on WTP, we need to control the three moving parts (i.e., WTP for a new item, disutility from a defect, and probability of a defect). We, therefore, designed framed lotteries with data from our pre-study (see Figure 1) that resemble the essential trade-offs consumers face when buying a remanufactured product. Next, we first explain the lotteries and then detail how framing is used to control the three moving parts.

Lotteries Subjects were asked to state their WTP for a lottery that probabilistically grants either a high outcome of 400 experimental units (EU) or a low outcome of 120 EU. The given probabilities were either exact (i.e., 25% chance of a low outcome) or ambiguous (probability ranges between 0% and 50%). As an example, a risk-neutral subject would be willing to pay $0.25 \cdot 120 + 0.75 \cdot 400 = 330$ EU (= WTP) for playing the lottery with an exact probability of 25% for the low outcome.

In order to provide incentives to report the WTP for the lottery truthfully, we employed the BDM method, as shown in Figure 2. All subjects were provided with an endowment of 400 EU and asked for their WTP. The amount to be paid for playing the lottery ("selling price") is determined by a random draw that follows a discrete uniform distribution with support between 0 and 400 EU. If the WTP was higher than the randomly chosen selling price, as shown in the example in Figure 3(a), then the subject had to pay the selling price for playing the lottery. On the other hand, if the announced WTP was lower than the randomly chosen price, then the



(a): Subject plays the lottery (b): Subject does not play the lottery

Figure 3: BDM method for framed lotteries: two examples of outcomes

subject did not play the lottery, and received a payoff of 400 EU (= endowment); see Figure 3(b) for an example.

Next, the experimenter determined the outcome of the lottery with a random draw from an urn. The urn contained 100 balls in two colors, each color associated with either the low outcome (120EU) or the high outcome (400EU). In the treatment with exact probabilities, the number of low-outcome colored balls was 25. In treatments with ambiguous probabilities, the fraction of each color was not known to the subjects; actually, the urn only contained balls of the color associated with the high outcome, as we justify below. In the event of playing the lottery, the subject's payoff was: Endowment (400 EU) minus (random) selling price plus the value of the product (120 EU for the low outcome, or 400 EU for the high outcome). In the example of Figure 3(a), the high outcome is realized.

Framing and Treatments We framed our lotteries as buying a remanufactured product instead of "buying" a lottery; see A. The low outcome of 120 EU refers to the case of receiving a remanufactured phone with quality issues, which matches the median WTP value announced by subjects in our MTurk study for products with known quality issues. The high outcome of 400 EU refers to buying a remanufactured phone that is as good as new. Because the high outcome of 400 EU corresponds to a "as good as new" phone, then this design choice controls for the WTP for new phones because a new phone is valued at 400 EU for all subjects. We opted for this design choice since Fox and Weber (2002) and Heath and Tversky (1991) show that framing matters. We, therefore, believe that framing gives a better proxy for the true WTP distribution of remanufactured phones. We further note that we are not aware of any risky-choice study that elicited WTP for lotteries (framed or unframed) that involved the probabilities that are relevant to the situation we are considering.

We further announced in the treatments with ambiguous probabilities (between 0 and 50% for a phone with quality issues) that the true probability of receiving a phone with quality issues is based on empirical estimates. Thus, subjects were required to form beliefs about quality issues in reality, which underlines the need to frame our lotteries. In reality, we applied a 0% probability to the low outcome, since zero cosmetic and functional defects resemble what remanufacturing tries to achieve.

Summarizing, we control the three moving parts that may otherwise confound with the effect of risk preferences on WTP: (1) We induce a WTP of 400 EU for a new phone by setting the high outcome of the lottery (as good as new) to 400 EU. (2) We induce a WTP of 120 EU for a phone with quality issues. (3) We control the risk perceptions by setting the exact (or ambiguous) probabilities of the lottery.

We have a total of four treatments in our experimental design, as shown in Table 3, where we manipulate the phone brand, and the probability of receiving a product with quality issues (i.e., the probability of a low outcome). In the iPhone treatment involving risk (iPR), we announce a probability of 25% for the low outcome. In the ambiguity treatments, we announce that the probability of a low outcome ranges between 0% and 50%, which approximately matches the range stated by the subjects in our MTurk survey. We have an ambiguity treatment for the iPhone (iPhone ambiguity, iPA) and another for the Huawei P9 (Huawei ambiguity, HA). Finally, we have a treatment where we cross-examine brand effects and the difference between purely risky and ambiguous lotteries (risk vs. ambiguity, RA). In the RA treatment, all participating subjects provide WTP for both phones, iPhone and Huawei, in a randomly assigned sequence, since our online survey results indicated that the presentation order was not important.

Note that there is always a combination of two treatments in which we manipulate exactly one factor. This allows us to establish root-cause effects. By comparing iPA against iPR, we can observe possible different effects of ambiguity and risk while holding the focal product constant. By comparing iPA with HA, we can observe possible brand effects while holding the framed lottery constant. Finally, by comparing iPR (alternatively, HA) with the WTP statement for the iPhone (alternatively, Huawei P9) in RA, we can observe the possible effect of comparing risky and ambiguous perceptions simultaneously (e.g., see literature review for a discussion of the comparative ignorance effect).

Table 3: Treatments

Treatment	Independent Observations	Probability of receiving a product with quality issues
iPhone ambiguity (iPA)	30	unknown 0%-50%
Huawei ambiguity (HA)	28	unknown 0%-50%
iPhone risk (iPR)	29	25%
Risk versus ambiguity (RA)	30	unknown 0%-50% for Huawei (RA-HA) 25% for iPhone (RA-iPR)

Protocol and Incentives The participants for the paper-pencil experiments were recruited online using ORSEE (Greiner, 2004). The subject pool contained graduate and undergraduate students of a mid-sized university in Germany. Each of the four treatments was administered in a single session that lasted less than 60 minutes. Upon arrival, participants received written instructions that were read out loud (see A). For each treatment we use a questionnaire in a between-subjects design, for a total of 117 participants; questions were answered privately, in the subject's cubicle at the lab.

Each subject, in all treatments, was required to complete two tasks: (i) a treatment-specific assessment of their WTP for the remanufactured product, which happened in a single round; and (ii) completion of the same survey as in the MTurk study. Subjects were only allowed to participate if three control questions were successfully answered. For details see A.

For the first task, we randomly chose one subject from each session to receive his/her respective payoff as monetary compensation. This design has no significant effect on behavior, compared to a design where all subjects receive their payoff (Trautmann et al., 2011; Harrison et al., 2007; Armantier, 2006). We converted the payoff into € at the exchange rate 2EU=€1. For completing the second task, all subjects received a flat payment of €5. The total average payment per subject was €12.04; this is considerably lower than the payoffs indicated in Figure 3 because only one subject in each session received his/her payoff for the first task.

3.3 Results of the Behavioral Experiments

Figure 4 displays the histograms, one for each treatment, of observed WTP. We used bin widths of 10 to account for the fact that most subjects reported their WTP in multiples of 5 or 10. We have also excluded observations in which the reported WTP was outside the feasible value

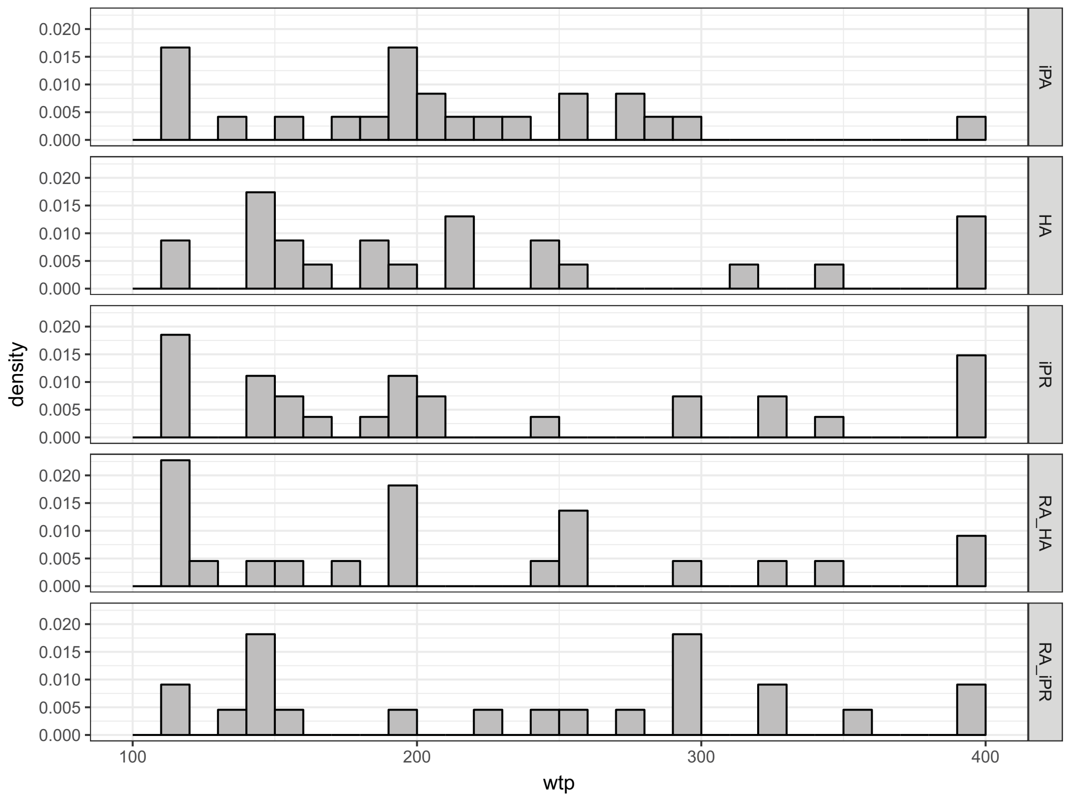


Figure 4: Histograms of WTP by treatments: iPA: iPhone ambiguity; HA: Huawei ambiguity; iPR: iPhone risk; RA-HA: risk vs. ambiguity with Huawei ambiguity; RA-iPR: risk vs. ambiguity with iPhone risk

range of $[120, 400]$.[2] As a result, we excluded 6, 5, 2, and 8 observations out of 30, 28, 29, and 30 observations from treatments iPA, HA, iPR, and RA, respectively; the final sample sizes are shown in Table 4.

Table 4 provides the averages and standard deviations of WTP for each of the four treatments. There is no significant treatment effect when comparing only the three treatments that include a single product, using a Mann-Whitney U-test: iPA vs. HA ($p < 0.99$), iPR vs. iPA ($p < 0.7$), HA vs. iPR ($p < 0.95$), where the p-values are two-sided.

We do find some evidence of the comparative ignorance effect when comparing the within-subject WTP statements in RA-HA vs. RA-iPR (Wilcoxon-test, two-sided, $p < 0.01$). In effect, participants who actively compare remanufactured devices with differing degrees of ambiguity tend to favor choices with less ambiguity. However, we do not find significant differences when comparing the respective single product treatments to the fourth treatment in which we simul-

[2] Reporting prices below 120 or above 400 cannot be captured by rational decision making and indicates that these participants did not understand the payoff structure. Such a boundedly rational behavior, however, may be captured by adding a noise term to the utility function in a discrete choice task (Luce, 1959); we have performed this analysis for robustness as discussed in section 4.

Table 4: Average WTP and standard deviation per treatment (see Table 2 for treatment definitions)

Treatment	iPA	HA	iPR	RA-HA	RA-iPR
Average	212.6	225.5	229.8	219.3	244.3
Std.dev.	68.3	90.2	99.8	91.7	91.7
Observations	24	23	27	22[a]	22[a]

[a]These 22 subjects completed both RA-HA and RA-iPR, as these were administered in a single session.

taneously compare a risky product to an ambiguous product: HA vs. RA-HA ($p < 0.66$) and iPR vs. RA-iPR ($p < 0.8$).

Given the observed variance in WTP (Figure 4), we next propose a model of consumer choice for remanufactured products that consistently explains the observed data. The choice model assumes that the observed heterogeneity in WTP is a result of heterogeneity in risk preferences. This assumption is supported by the empirical study from Abbey et al. (2017) who show that consumers' WTP is significantly linked to perceived quality risks even after controlling for other variables such as consumer traits (e.g., green traits, demographics, etc.).

4 Utility Model Capturing Heterogeneous Risk and Ambiguity Preferences

Given our empirical results, and the theoretical grounding in the literature discussed in sections 1 and 2 regarding risk and ambiguity, we propose the use of a consumer choice model from the risky-choice literature for use with remanufactured products. If such a theoretical model provides a good fit to our empirical results, then it can be used in analytical models that model consumer behavior for remanufactured products. Our proposed consumer choice model includes three facets of preferences: risk aversion, loss aversion, and ambiguity preferences. We discuss the importance of each of the three facets below. Risk and loss aversion are considered by the utility function from von Gaudecker et al. (2011), which assumes that subjects have a utility $U(x)$ as a function of the monetary payoff x, given by:

$$U(x) = \begin{cases} -\frac{1}{\gamma}e^{-\gamma(x-r)}, & \text{for } x \geq r \\ \frac{\lambda-1}{\gamma} - \frac{\lambda}{\gamma}e^{-\gamma(x-r)}, & \text{for } x < r. \end{cases} \tag{1}$$

In (1), $\lambda > 0$ is the loss aversion coefficient. For $\lambda > 1$, losses result in a larger disutility than gains of the same amount. A loss is perceived if the payoff x is smaller than a reference point r.

We consider reference points that are fixed (e.g., zero or the endowment of 400) or dependent on the WTP (i.e., $r(WTP)$, see Uppari and Hasija (2017) for a discussion of reference points in the newsvendor model). Further, $\gamma \in \Re$ is the Arrow-Pratt coefficient of absolute risk aversion where $\gamma < 0$ depicts risk-seeking behavior, $\gamma = 0$ depicts risk neutral preferences, and $\gamma > 0$ depicts risk aversion.

In a treatment involving only risk (there is no ambiguity), the expected utility of a consumer that reports his/her willingness to pay to be WTP is given by:

$$EU_{\text{risk}}(WTP) = \sum_{i=0}^{WTP} \frac{1}{401}[0.75 \cdot U(800-i) + 0.25 \cdot U(520-i)] + \left(1 - \frac{WTP+1}{401}\right) U(400). \quad (2)$$

In (2), the first term corresponds to the instances where there is a purchase (the randomly drawn price i is lower than WTP): with 75% probability, the product is as good as new so that the subject's payoff is the endowment of 400 plus the buy-back price of 400 minus the purchase price i; with 25% probability, the product is defective and the subject's payoff is the endowment of 400 plus the buy-back price of 120 minus the purchase price i. The second term in (2) corresponds to the instances where the randomly drawn price i is larger than the subject's reported WTP. A subject who maximizes the expected utility in (2) would report a WTP^* for which it holds that $EU_{risk}(WTP^*) - EU_{risk}(WTP^* - 1) \geq 0$ and $EU_{risk}(WTP^* + 1) - EU_{risk}(WTP^*) < 0$. This is equivalent to $0.75 \cdot U(800 - WTP^*) + 0.25 \cdot U(520 - WTP^*) \geq U(400) > 0.75 \cdot U(800 - (WTP^*+1)) + 0.25 \cdot U(520 - (WTP^*+1))$, that is, WTP^* is also the price at which the consumer is either indifferent between buying and not buying the product, or slightly prefers to buy the product (an effect that vanishes for continuous price distributions).

In our experiments, all subjects were provided with an endowment of 400 EU. As a result, losses (i.e., a negative payoff x in (1)) cannot occur even if the remanufactured product attains a low value. This corresponds to a setting in which the fixed reference point equals zero (i.e., $r = 0$), which is denoted as risk aversion only (RAO) model. Previous research, however, indicated that subjects perceive losses in lotteries with endowments because the endowment serves as the reference point for a subject's perception of gains and losses (see, e.g., Starmer and Sugden, 1991). Consequently, when one refines the consumer choice model by incorporating loss aversion in a treatment that only involves risk, a subject's fixed reference point becomes the endowment, i.e., $r = 400$ (risk aversion with loss aversion and a fixed reference point of $r = 400$; denoted as RALA). A discussion of reference points that depend on WTP, $r(WTP)$, is presented in D.

We now turn to the treatments with ambiguity. We model ambiguity aversion with the α-MEU ambiguity heuristic (Marinacci, 2002; Ghirardato et al., 2004), which has been reported to provide a good fit in decision tasks under ambiguity (Ahn et al., 2014). In this heuristic, it is assumed that subjects compare uncertain alternatives by putting a weight α on the minimum expected utility and a weight $(1-\alpha)$ on the maximum expected utility. Thus, a person appears to be ambiguity seeking (ambiguity averse) for $\alpha < 0.5$ ($\alpha > 0.5$). In the treatments with ambiguity, the minimum expected utility occurs when the probability of a defective product is 50%, whereas the maximum expected utility occurs when the probability of a defective product is zero. As a result, we obtain for our ambiguity treatment[3]:

$$
\begin{aligned}
EU_{\text{ambiguity}}(WTP) = \alpha\left(\sum_{i=0}^{WTP} \frac{1}{401}[0.5\cdot U(800-i) + 0.5\cdot U(520-i)]\right) + \qquad (3)\\
+ (1-\alpha)\left(\sum_{i=0}^{WTP} \frac{1}{401}[1\cdot U(800-i) + 0\cdot U(520-i)]\right) + \left(1-\frac{WTP+1}{401}\right)U(400),
\end{aligned}
$$

We use the utility functions (1)-(3) with probability distributions for the coefficients of loss aversion λ and risk aversion γ from the large-scale study conducted by von Gaudecker et al. (2011), which they find to be lognormal and normal, respectively. These estimates are shown in Table 5. The study by von Gaudecker et al. (2011) was carefully conducted with 1,422 subjects in Europe and thus should reasonably reflect the distribution of these parameters in the general population. However, for robustness purposes, we have performed our own estimation of the *mean* value of the same parameters with our sample by using a maximum likelihood estimation (MLE) procedure in a random utility model. Our results, despite the relatively small sample size, are reasonably consistent. Specifically, our estimates of γ and λ are 0.028 and 1, compared to the mean values of 0.03 and 0.9 in von Gaudecker et al. (2011); please see C for details. We use the results from von Gaudecker et al. (2011) due to their much larger sample size in addition to their estimation of the standard deviation of such parameters.

With respect to α, we are not aware of a study that estimates **all three** parameters $(\alpha, \lambda, \gamma)$ simultaneously. Indeed, we were also unsuccessful in estimating the mean value of the three parameters simultaneously using MLE, as detailed in C. As a result, we employ the normal distribution estimate for the ambiguity parameter α from the only study that, to our knowledge, estimates the parameter: Ahn et al. (2014) as shown in Table 5. The mean value of α is 0.52, which means that, on average, there is little ambiguity aversion in the general population, at least in the context of Ahn et al. (2014). However, as discussed in Section 2, context matters for ambiguity. Thus, for robustness we also consider two additional uniform distributions for α, with a similar variability to Ahn et al. (2014), but larger means of 0.6 and 0.7. In other

[3]Note that equation (3) nests equation (2) for $\alpha = 0.5$. Also, in this case, the utility maximizing WTP^* reflects a price at which the consumer is essentially indifferent between buying and not buying the product.

Table 5: Parameter used in the utility functions

Symbol	Definition	μ	σ
γ	Coefficient of risk aversion	0.03	0.04
λ	Coefficient of loss aversion	0.90	1.53
α (from Ahn et al., 2014)	Ambiguity parameter	0.52	0.12

words, the robustness examination assesses a population that is, on average, ambiguity averse: $\alpha \sim U[0.4, 0.8]$ and $\alpha \sim U[0.5, 0.9]$. Note that when $\alpha \sim U[0.5, 0.9]$, all consumers are ambiguity averse.[4]

For a given combination of λ, γ, and α, we can calculate numerically, through a line search, the WTP (in €) that maximizes expected utility in (2) and (3), which, as previously discussed, is the same value of WTP at which the consumer is indifferent between buying and not buying the product. Given the probability distributions for the parameters λ, γ, and α given in Table 5, we can thus estimate a theoretical probability distribution of WTP, for a given treatment, through Monte-Carlo simulation. To that end, we generated 10,000 simulated consumers by independently drawing random values for each parameter. This was coded in MATLAB.

Results Table 6 displays the mean and standard deviation of the distribution of WTP for the different theoretical consumer choice models, including different combinations of loss and ambiguity aversion as summarized in the table. We now compare these different theoretical consumer choice models against the empirically observed WTP distributions. The results for WTP-dependent reference points—$r(WTP)$—are provided in D.

Figure 5 shows the empirical WTP histograms for each of the treatments, where, contrary to Figure 4, the bin size is set to 40 for better visualization, along with two curves. The solid line represents a fitted smooth kernel density of the observed data—the best fitting smooth empirical kernel density. The dotted line represents the density of the theoretical consumer choice model that best fits the data—the RAO model from Table 6, which incorporates only risk aversion. (Plots of the theoretical WTP distributions for the other consumer choice models from Table 6 are provided in B.) The best-fit conclusion originates from χ^2-tests of the different theoretical WTP distributions against the empirical distributions shown in Figure 4. We use the same sub-categories of width 10 but combine sub-categories such that the expected frequency in a test category exceeds 5 (Sheskin, 2004). The critical values for a 10% level of significance with

[4] Another robustness test with considerably larger variability, $\alpha \sim U[0, 1]$, showed similar results, and is therefore not reported here.

Table 6: Mean WTP and standard deviation per model

Theoretical model	Loss aversion?	Ambiguity aversion?	μ	σ
Risk aversion				
- only (RAO)	no	no	234.5	78.2
- with loss aversion, $r = 400$ (RALA)	yes	no	206.6	82.8
Ambiguity aversion (α distribution according to Ahn et al., 2014)				
- with risk aversion only (AARA)	no	yes	234.0	78.8
- with loss aversion, $r = 400$ (AALA)	yes	yes	206.4	83.0
Ambiguity aversion ($\alpha \sim U[0.4, 0.8]$; for robustness check)				
- with risk aversion only (AARA)	no	yes	225.9	77.6
- with loss aversion, $r = 400$ (AALA)	yes	yes	200.0	81.0
Ambiguity aversion ($\alpha \sim U[0.5, 0.9]$; for robustness check)				
- with risk aversion only (AARA)	no	yes	217.3	75.9
- with loss aversion, $r = 400$ (AALA)	yes	yes	193.5	78.7

four and three degrees of freedom are $\chi^2_{4,0.1} = 7.78$, and $\chi^2_{3,0.1} = 6.25$, respectively. Table 7 summarizes the test statistics. In Table 7, numbers in bold indicate that we cannot reject the null hypothesis that the empirical results can be explained by the respective utility model.[5]

Somewhat surprisingly, we find that the simplest theoretical model that includes only risk aversion (RAO) organizes the observations most consistently. The null hypothesis is not rejected for any of the treatments under the model RAO, while all other models are rejected for at least one of the treatments. While including loss aversion (RALA) improves the fit in the iPR treatment, it performs worse in the other treatments. Similarly, the models including ambiguity (AA*) provide a slightly better fit for the iPA treatment but are worse for the others. This also holds true for the treatments in which the ambiguous option and the unambiguous option are compared (RA-**). While this may be surprising in light of our within-subject effect on WTP (RA-HA vs. RA-iPR, Wilcoxon-test, two-sided, $p < 0.01$, discussed before), we find no significant effect in comparison to the single product treatments (HA vs. RA-HA ($p < 0.66$) and iPR vs. RA-iPR ($p < 0.8$)). In sum, we find that the more parsimonious model that includes only risk aversion is a reasonable and consistent approximation of consumer choice behavior.

[5]For robustness, we have also performed Kolmogorov-Smirnov (K-S) tests based on sub-categorized data. The tests confirm the results from the χ^2-tests.

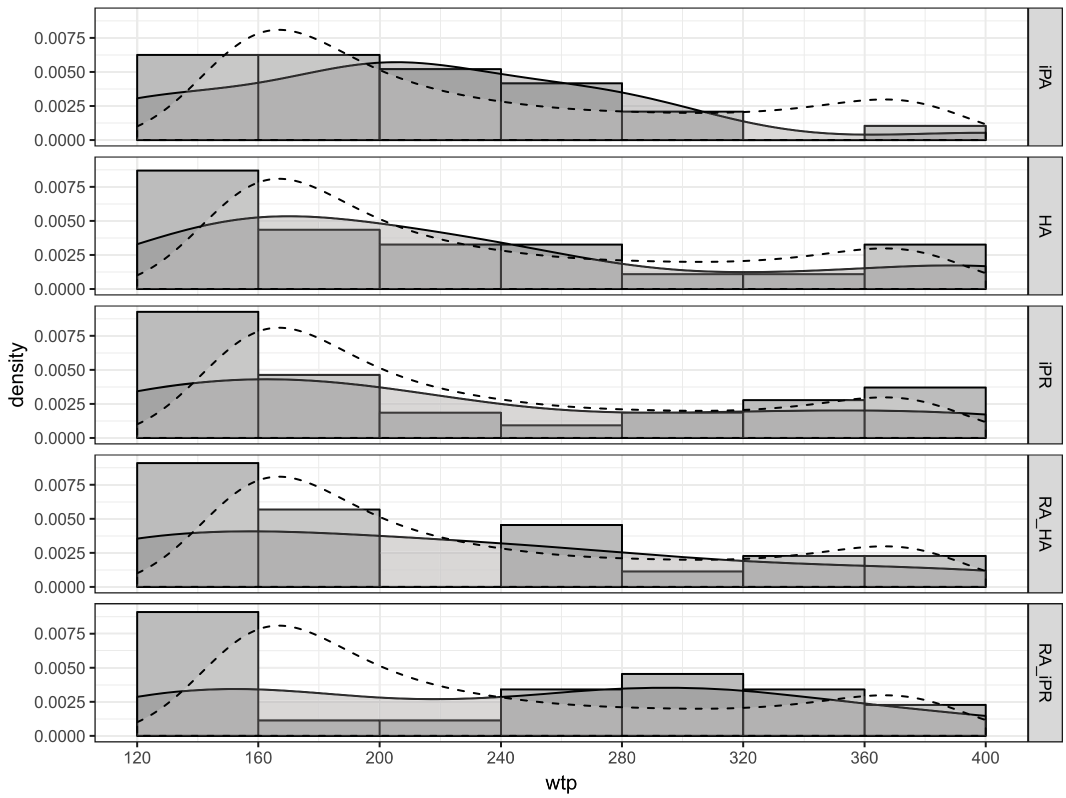


Figure 5: Histograms (bin size equal to 40) of observed data in each of the treatments. The solid lines represent kernel density estimates of the observed data. The dotted lines represent the theoretical consumer choice distribution for the risk-only (RAO) model

5 Conclusion

The CLSC literature makes it clear that consumers' WTP are lower for remanufactured products than for corresponding new ones. This relationship is at the heart of many CLSC models that provide insights for managers regarding the introduction, pricing, and marketing of remanufactured products, although they assume constant discount factors for remanufactured products among the population. Abbey et al. (2017) show that disregarding the heterogeneity in discount factors results in an under-pricing of remanufactured products and an understatement of the potential profit for an OEM when entering this market. They obtain these insights using an empirically derived WTP distribution. Our work provides further evidence that a relatively simple utility model that incorporates only risk aversion, with known estimates of the risk aversion distribution in the general population, is a reasonable refinement of the constant discount factor model thus far used in the CLSC literature. This model refinement will hopefully inspire researchers to derive continued insights without the burden of empirically deriving WTP distributions. Specifically, our results were obtained using known risk preferences estimates from the general population, not from an over-fitting of consumer choice models to our data.

Table 7: Test statistics of the χ^2-test (degrees of freedom in brackets, if different from 3). Bold numbers indicate that the null hypothesis is **not** rejected at a 10% level of significance. The smallest row value is underlined. Greyed-out values indicate that ambiguity should not matter in treatments involving risk only.

	Theoretical model				Uniform α (robustness checks)			
					$\alpha \sim U[0.4, 0.8]$		$\alpha \sim U[0.5, 0.9]$	
Treatment	RAO	RALA	AARA	AALA	AARA	AALA	AARA	AALA
iPA	**5.81**	10.74	**<u>5.42</u>**	10.50	9.10	6.77	6.70	8.84
HA	**1.51**	**6.47**	**3.69**	6.51	**<u>1.15</u>**	**2.93**	**3.04**	**4.38**
iPR	**5.92**	**3.59** (4)	**5.12**	**3.51** (4)	**3.29**	**4.10**	**<u>3.17</u>** (4)	**5.85**
RA-HA	**<u>0.83</u>**	**2.42**	8.23	**2.28**	**1.37**	**3.01**	**4.42**	**3.70**
RA-iPR	**<u>5.44</u>**	**6.09**	9.05	6.32	7.88	9.70	9.83	11.82

In addition, our results provide new insights on the behavioral factors that drive WTP for remanufactured products: ambiguity and loss aversion only have a minor effect, while risk aversion has a major impact on WTP. Enabling consumers to resolve the ambiguity related to the remanufacturing process alone is likely not to have a major impact on the overall WTP distribution. Instead, firms should focus on reducing the perceived risk of remanufactured products, perhaps by providing information to consumers on the results of tests of performance against corresponding new products.

As with any work, our study has limitations. First, our results only considered one product type for consistency and comparability with earlier studies (e.g., Agrawal et al. (2015); Abbey et al. (2017)). The results from Abbey et al. (2017), who found their iPhone results to be robust to other Apple products such as the iPod and the MacBook Air, suggest that our model may hold for a wider range of consumer electronics products. Yet, it is not clear if the results will be applicable to other brands in their respective product categories, such as appliances or mechanical products.

Second, our laboratory experiments allow us to cleanly focus on the risky-choice aspects of the buying decision, thereby enabling us to derive cause-effect conclusions regarding the impact of risk, ambiguity and loss aversion with high internal validity. However, the requirement to control for the moving parts that determine WTP (i.e., WTP for new products, probability of/disutility from defects) restricts us to elicit WTP for framed lotteries, as opposed to the direct elicitation of WTP for the products. While there are other methods to directly elicit WTP for new or remanufactured products (e.g., see the incentive-aligned conjoint framework in Ding, 2007; Agrawal et al., 2015) by physically providing the product at the end of the experiment, these methods make it essentially unfeasible to induce/control the probability of/disutility from defects, which is essential for linking heterogeneity in WTP to idiosyncratic risk preferences. We note, however, that experiments in natural environments limit the ability to control the theory's moving parts (see Katok, 2011). Nonetheless, empirical triangulation through other methods may continue to strengthen the external validity of our results and those of other related studies.

References

Abbey, J.D., Blackburn, J.D., Guide, V.D.R., 2015a. Optimal pricing for new and remanufactured products. Journal of Operations Management 36, 130–146.

Abbey, J.D., Kleber, R., Souza, G.C., Voigt, G., 2017. The role of perceived quality risk in pricing remanufactured products. Production and Operations Management 26, 100–115.

Abbey, J.D., Meloy, M., V. Daniel Guide, J., Atalay, S., 2015b. Remanufactured products in closed-loop supply chains for consumer goods. Production and Operations Management 24, 488–503.

Abdellaoui, M., Baillon, A., Placido, L., Wakker, P.P., 2011. The rich domain of uncertainty: Source functions and their experimental implementation. The American Economic Review 101, 695–723.

Agrawal, V., Atasu, A., Ittersum, K.V., 2015. Remanufacturing, third-party competition, and consumers' perceived value of new products. Management Science 61, 60–72.

Ahn, D., Choi, S., Gale, D., Kariv, S., 2014. Estimating Ambiguity Aversion in a Portfolio Choice Experiment. Quantitative Economics 5, 195–223.

Akturk, M., Abbey, J., Geismar, H., 2017. Strategic design of multiple lifecycle products for remanufacturing operations. IISE Transactions 49, 967–979.

ANSI Consensus Body, 2015. Remanufacturing industries council consensus body to form an American National Standards Institute (ANSI) standard for remanufacturing. Conditionally Approved Draft Standard ANSI/RIC: 2015 9-22-2015.

Armantier, O., 2006. Do wealth differences affect fairness considerations? International Economic Review 47, 391–429.

Atasu, A., Sarvary, M., van Wassenhove, L., 2008. Remanufacturing as a marketing strategy. Management Science 47, 1081–1100.

Atasu, A., Souza, G.C., 2013. How does product recovery affect quality choice? Production and Operations Management 22, 991–1010.

Atasu, A., V. Daniel Guide, J., Van Wassenhove, L., 2010. So what if remanufacturing cannibalizes my new product sales? California Management Review 52, 56–76.

Becker, G.M., Degroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. Systems Research 9, 226–232.

Camerer, C., 1995. Individual decision making, in: Kagel, J.H., Roth, A.E. (Eds.), The handbook of experimental economics. Princeton University Press, Princeton, NJ, pp. 587–703.

Camerer, C., Weber, M., 1992. Recent developments in modeling preference: Uncertainty and ambiguity. Journal of Risk and Uncertainty 5, 325–370.

Charness, G., Gneezy, U., Imas, A., 2013. Experimental methods: Eliciting risk preferences. Journal of Economic Behavior and Organization 87, 43–51.

Chow, C.C., Sarin, R.K., 2001. Comparative ignorance and the ellsberg paradox. Journal of Risk and Uncertainty 22, 129–139.

Cohen, M., Jaffray, J.Y., Said, T., 1987. Experimental comparison of individual behavior under risk and under uncertainty for gains and for losses. Organizational Behavior and Human Decision Processes 39, 1–22.

Curley, S.P., Yates, J.F., 1989. An empirical evaluation of descriptive models of ambiguity reactions in choice situations. Journal of Mathematical Psychology 33, 397–427.

Debo, L.G., Toktay, L.B., Van Wassenhove, L.N., 2005. Market segmentation and product technology selection for remanufacturable products. Management Science 51, 1193–1205.

Ding, M., 2007. An incentive-aligned mechanism for conjoint analysis. Journal of Marketing Research 44, 214–223.

Einhorn, H.J., Hogarth, R.M., 1988. Decision making under ambiguity: A note, in: Munier, B.R. (Ed.), Risk, Decision and Rationality. Springer Netherlands, Dordrecht, pp. 327–336.

Ellsberg, D., 1961. Risk, ambiguity, and the Savage axioms. The Quarterly Journal of Economics 75, 643–669.

Etner, J., Jeleva, M., Tallon, J.M., 2012. Decision theory under ambiguity. Journal of Economic Surveys 26, 234–270.

Ferguson, M., Guide, Jr., V.D.R., Souza, G.C., 2006. Supply chain coordination for false failure returns. Manufacturing & Service Operations Management 8, 376–393.

Ferguson, M., Souza, G.C., 2010. Closed-Loop Supply Chains: New Developments to Improve the Sustainability of Business Practices. CRC Press, Boca Raton, FL.

Ferguson, M.E., Toktay, L.B., 2006. The effect of competition on recovery strategies. Production and Operations Management 15, 351–368.

Ferrer, G., Swaminathan, J.M., 2006. Managing new and remanufactured products. Management Science 52, 15–26.

Fox, C.R., Tversky, A., 1995. Ambiguity aversion and comparative ignorance. The Quarterly Journal of Economics 110, 585–603.

Fox, C.R., Tversky, A., 1998. A belief-based account of decision under uncertainty. Management Science 44, 879–895.

Fox, C.R., Weber, M., 2002. Ambiguity aversion, comparative ignorance, and decision context. Organizational Behavior and Human Decision Processes 88, 476–498.

Galbreth, M.R., Boyaci, T., Verter, V., 2013. Product reuse in innovative industries. Production and Operations Management 22, 1011–1033.

von Gaudecker, H.M., van Soest, A., Wengstrom, E., 2011. Heterogeneity in risky choice behavior in a broad population. The American Economic Review 101, 664–694.

Ghirardato, P., Maccheroni, F., Marinacci, M., 2004. Differenciating ambiguity and ambiguity attitude. Journal of Economic Theory 118, 133–173.

Ghosh, D., Ray, M.R., 1997. Risk, ambiguity, and decision choice: Some additional evidence. Decision Sciences 28, 81–104.

Gilboa, I., Schmeidler, D., 1989. Maxmin expected utility with non-unique prior. Journal of Mathematical Economics 18, 141–153.

Gneezy, U., Imas, A., List, J., 2015. Estimating individual ambiguity aversion: A simple approach. Technical Report. National Bureau of Economic Research.

Greiner, B., 2004. An online recruitment system for economic experiments, in: Kremer, K., Macho, V. (Eds.), Forschung und wissenschaftliches Rechnen 2003, GWDG Bericht 63. Gesellschaft für Wissenschaftliche Datenverarbeitung, Göttingen, pp. 79–93.

Guide, Jr., V.D.R., Li, K., 2010. The potential for cannibalization of new product sales by remanufactured products. Decision Sciences 41, 547–572.

Halevy, Y., 2007. Ellsberg revisited: An experimental study. Econometrica 75, 503–536.

Harrison, G.W., Lau, M.I., Rutström, E.E., 2007. Estimating risk attitudes in denmark: A field experiment. Scandinavian Journal of Economics 109, 341–368.

Heath, C., Tversky, A., 1991. Preference and belief: Ambiguity and competence in choice under uncertainty. Journal of Risk and Uncertainty 4, 5–28.

Kahneman, D., Knetsch, J.L., Thaler, R.H., 1990. Experimental tests of the endowment effect and the coase theorem. Journal of Political Economy 98, 1325–1348.

Katok, E., 2011. Using laboratory experiments to build better operations management models. Foundations and Trends in Technology, Information and Operations Management 5, 1—86.

Kőszegi, B., Rabin, M., 2006. A model of reference-dependent preferences. The Quarterly Journal of Economics 121, 1133–1165.

Köbberling, V., Wakker, P., 2005. An index of loss aversion. Journal of Economic Theory 122, 119–131.

Long, X., Nasiry, J., 2014. Prospect theory explains newsvendor behavior: The role of reference points. Management Science 61, 3009–3012.

Luce, R.D., 1959. Individual Choice Behavior: A Theoretical Analysis. Wiley, New York.

Lund, R.T., 1984. Remanufacturing. Technology Review 87, 19–27.

Marinacci, M., 2002. Probabilistic sophistication and multiple priors. Econometrica 70, 755–764.

Oraiopoulos, N., Ferguson, M., Toktay, L.B., 2012. Relicensing as a secondary market strategy. Management Science 58, 1022–1037.

Örsdemir, A., Kemahlioğlu-Ziya, E., Parlaktürk, A.K., 2014. Competitive quality choice and remanufacturing. Production and Operations Management 23, 48–64.

Ovchinnikov, A., 2011. Revenue and cost management for remanufactured products. Production and Operations Management 20, 824–840.

Ovchinnikov, A., Blass, V., Raz, G., 2014. Economic and environmental assessment of remanufacturing strategies for product + service firms. Production and Operations Management 23, 744–761.

Schmeidler, D., 1983. Subjective probability and expected utility without additivity. Econometrica 51, 571–587.

Sheskin, D., 2004. Handbook of parametric and nonparametric statistical procedures. Third ed., Chapman & Hall/CRC, Boca Raton.

Smith, K., Dickhaut, J., McCabe, K., Pardo, J.V., 2002. Neuronal substrates for choice under ambiguity, risk, gains, and losses. Management Science 48, 711–718.

Souza, G.C., 2013. Closed-loop supply chains: A critical review, and future research. Decision Sciences 44, 7–38.

Starmer, C., 2000. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. Journal of Economic Literature 38, 332–382.

Starmer, C., Sugden, R., 1991. Does the random-lottery incentive system elicit true preferences? An experimental investigation. The American Economic Review 81, 971–978.

Subramanian, R., Ferguson, M., Toktay, L.B., 2013. Remanufacturing and the component commonality decision. Production and Operations Management 22, 36–53.

Subramanian, R., Subramanyam, R., 2012. Key factors in the market for remanufactured products. Manufacturing & Service Operations Management 14, 315–326.

Trautmann, S.T., Schmidt, U., 2012. Pricing risk and ambiguity: the effect of perspective taking. Quarterly Journal of Experimental Psychology 65, 195–205.

Trautmann, S.T., Vieider, F.M., Wakker, P.P., 2011. Preference reversals for ambiguity aversion. Management Science 57, 1320–1333.

Tversky, A., Kahneman, D., 1991. Loss aversion in riskless choice: A reference-dependent model. The Quarterly Journal of Economics 106, 1039–1061.

Uppari, B.S., Hasija, S., 2017. Modeling newsvendor behavior: A prospect theory approach. INSEAD Working Paper No. 2017/66/DSC/TOM. Available at SSRN: https://ssrn.com/abstract=2485149.

Von Neumann, J., Morgenstern, O., 1947. Theory of games and economic behavior, 2nd ed. Princeton University Press.

Vorasayan, J., Ryan, S.M., 2006. Optimal price and quantity of refurbished products. Production and Operations Management 15, 269–383.

A Questionnaire used in experiments

Instructions

Decision Situation:

Please, state your maximum willingness to pay for the remanufactured iPhone SE shown below.

iPhone SE – remanufactured
(Retail New $399)

Technical Details
- Capacity: 32 GB
- Battery: Up to 73 hours of use
- Screen: HD Resolution
- Camera: 12 megapixel camera with *f*/2.2 aperture and 1080p HD video recording
- Security: Fingerprint identity sensor
- Warranty: Same as new

The remanufactured iPhone SE has with an unknown probability between **0% and 50%** a cosmetic defect. In this case, the product has a value of **120 EU**. Otherwise, the product is as-good-as new and has a value of **400 EU**. What is your maximum willingness to pay for this remanufactured iPhone SE? In other words, how much are you willing to pay for this remanufactured iPhone SE, without knowing whether a cosmetic defect is present?

Compensation:

At the end of the experiment, **one** participant will be randomly chosen and compensation will be paid. Every participant has the same chance of being chosen. The other participants receive no compensation out of this decision situation but they receive a compensation when completing the survey in the second part of the experiment.

You have **400 EU to start with**.

There are 401 balls in a bag, which are labeled from 0 to 400. The experimenter will draw a ball at the end of the experiment, which represents the price of the remanufactured phone. If the price (=number on the ball) is larger than your willingness to pay, you receive 400 EU. If the price is larger than your willingness to pay, you will receive the initial 400 EU. If the price is smaller or equal to your willingness to pay, the initial capital is reduced by the price and you receive the value of the remanufactured iPhone SE, which is determined as follows:

The value will be determined by drawing a ball from an Urn that contains in total 100 balls that are either red or black. Drawing a red ball gives you 120 EU (the equivalent to the value of a reman with a cosmetic defect). Drawing a black ball gives you 400 EU (the equivalent to the value of a reman that is “as-good-as new”). The number of red balls (“cosmetic defect”) was determined based on empirical data on remanufactured products. The exact number of red balls is unknown to you and ranges between 0 and 50. In turn, the number of black balls (“as-good-as new”) may range – based on empirical data - between 50 and 100.

How to determine my payoff?

Your payoff is determined as follows:

a) Willingness to pay is larger or equal to selling price:

   Payoff: starting capital – selling price + value of the iPhone

b) Willingness to pay is smaller than selling price:

   Payoff: starting capital

Finally, the payoff is converted into € at an exchange rate: 2EU = 1€

Control question 1: You have stated a willingness to pay of 100EU. The random price is 50EU. A red ball (“cosmetic defect”) is drawn. What will be the payoff in EU?

Control question 2: You have stated a willingness to pay of 100EU. The random price is 550EU. A red ball (“cosmetic defect”) is drawn. What will be the payoff in EU?

Control question 3: You have stated a willingness to pay of 100EU. The random price is 50EU. A black ball (“as-good-as new”) is drawn. What will be the payoff in EU?

Please state your maximum willingness to pay!

My willingness to pay (in EU) is:

B Plots of kernel densities for other consumer choice models

Figure 6 shows theoretical WTP histograms along with smooth kernel densities for all six consumer choice models, including different risk, loss, and ambiguity combinations.

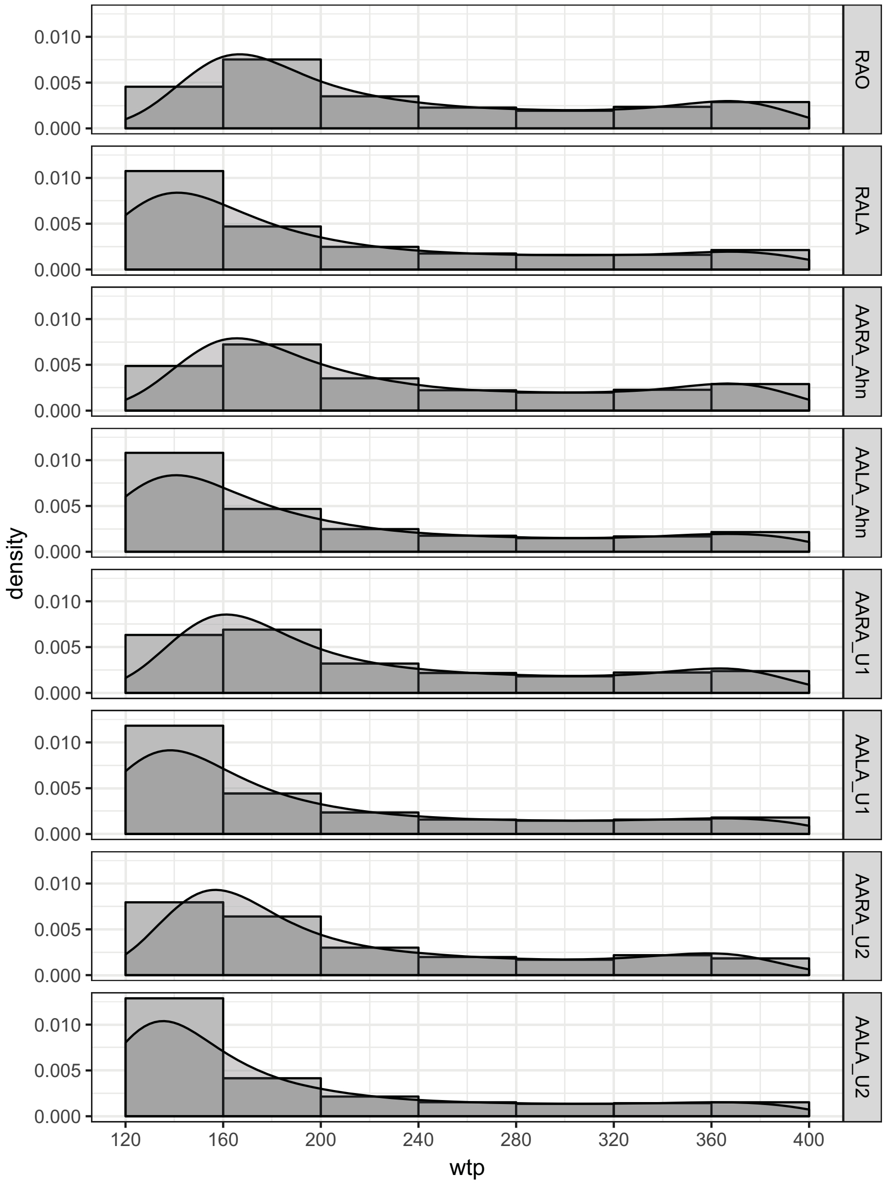


Figure 6: Histograms (bin size equal to 40) and kernel densities estimates of willingness to pay (WTP) by utility function (U1 is for $\alpha \sim U[0.4, 0.8]$ and U2 is for $\alpha \sim U[0.5, 0.9]$).

C MLE estimation of risk, loss, and ambiguity parameters

In this section, we detail the procedure for estimating the risk, loss and ambiguity parameters using a maximum likelihood estimation (MLE) procedure, as a robustness test. We rely on a random utility model: $U_i = EU(\alpha, \lambda, \gamma, WTP_i) + \epsilon_i$, where EU is given in (2)-(3), and ϵ_i follows

Table 8: MLE estimation with $\theta = 0.1$

Parameter	Null model	Risk only	Risk and loss	Risk, loss, and ambiguity
θ	0.01	0.01	0.01	0.01[a]
γ	-	0.028	0.028	-[a]
λ	-	-	1	-[a]
α	-	-	-	-[a]
log-L	-7674.4	-655.6	-655.6	-[a]

[a]We were not able to obtain reliable estimates for this model.

an extreme value distribution and depicts the unobserved variance. Under such assumptions, the choice probability for a stated $WTP_i \in \{120, 130, ..., 400\}$ is:

$$\Pr(WTP_i) = \frac{e^{\theta \cdot U_i}}{\sum_{WTP_j \in \{120,130,...,400\}} e^{\theta \cdot U_j}}. \tag{4}$$

Here, θ measures how much weight the utility function has in explaining the choice behavior. As $\theta \to 0$, all exponentials become one, and each WTP_i value is chosen with equal probability. The likelihood function becomes $L(\alpha, \lambda, \gamma, \theta) = \Pi_{j=1}^{N} \Pr(WTP_j)$, where N is the number of observations. Note that we only estimate α for those treatments in which we induce ambiguity. We maximized the log-likelihood function using the `maxLik` package in R, with the Nelder-Mead optimization routine and an iteration limit of 100. For all models (risk only, risk and loss, risk, loss and ambiguity), we obtain an estimate of θ near zero, which means that all WTP_i options are chosen with nearly equal probability. That is, the behavioral model does not explain the WTP observations. We believe that this is due to the large number of parameters to be estimated (four) with a small sample size around 100 observations.

As a result, we have attempted to estimate $(\alpha, \lambda, \gamma)$ for *given* values of θ. After several trials, it was possible to obtain meaningful estimates of α and λ, but not of α, as shown in Table 8. Note that the estimate of γ for the models with risk only, and risk and loss aversion is 0.028, which is virtually identical to the mean value $\gamma = 0.03$ from von Gaudecker et al. (2011), obtained through a completely different estimation method. In addition, the estimated value of λ in the risk and loss aversion model is 1, which means no loss aversion; the mean value reported in von Gaudecker et al. (2011) is 0.9, indicating that, on average, the loss aversion effect in the population is rather small. Our results for this small sample are thus reasonably consistent with those of von Gaudecker et al. (2011). We were not able to estimate the parameter α simultaneously with γ and λ in the model with risk, loss and ambiguity aversion; this is again due to the small sample size, as then three parameters need to be estimated.

D Robustness test: WTP-dependent reference points

In line with Long and Nasiry (2014) and Uppari and Hasija (2017), WTP-dependent reference points, $r(WTP)$, have been obtained as a convex combination of the maximum possible payoff (800) and the minimum possible payoff, $\min(400, 520 - WTP)$, as follows

$$r(WTP) = \beta \cdot 800 + (1 - \beta) \cdot \min(400, 520 - WTP) \quad (5)$$

In (5), β denotes the decision maker's level of optimism. The test statistics provided in Table 9 quickly degrade with decreasing level of optimism. Please note that in the case of $\beta = 1$, utilities all fall into the loss domain. Thus, loss aversion does not impact WTP, yielding the same results as the risk aversion only (RAO) model. Table 9 shows that the RAO model has the best fit in all treatments.

Table 9: Test statistics of the χ^2-test (degrees of freedom in brackets, if different from 3). Bold numbers indicate that the null hypothesis is **not** rejected at a 10% level of significance.

	Theoretical model	Dynamic reference points			
Treatment	RAO	$\beta = 1$	$\beta = 0.75$	$\beta = 0.5$	$\beta = 0.25$
iPA	**5.81**	**5.81**	22.81	28.73	22.14 (2)
HA	**1.51**	**1.51**	12.07	19.45	12.46 (2)
iPR	**5.92**	**5.92**	19.39	35.87	9.53 (2)
RA-HA	**0.83**	**0.83**	22.38	16.03	8.56 (2)
RA-iPR	**5.44**	**5.44**	9.81	9.79	7.03 (2)

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

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

www.fww.ovgu.de/femm

ISSN 1615-4274