



Universität Hamburg  
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Fakultät Wirtschafts- und  
Sozialwissenschaften



# Testing the Endowment Effect for Default Rules

Isabel Marcin  
Andreas Nicklisch

**WiSo-HH Working Paper Series**  
Working Paper No. 10  
January 2014



WiSo-HH Working Paper Series  
Working Paper No. 10  
January 2014

## Testing the Endowment Effect for Default Rules

Isabel Marcin, Max Planck Institute for Research on Collective Goods, Bonn  
Andreas Nicklisch, Universität Hamburg

ISSN 2196-8128

Font used: „TheSans UHH“ / LucasFonts

Die Working Paper Series bieten Forscherinnen und Forschern, die an Projekten in Federführung oder mit der Beteiligung der Fakultät Wirtschafts- und Sozialwissenschaften der Universität Hamburg tätig sind, die Möglichkeit zur digitalen Publikation ihrer Forschungsergebnisse. Die Reihe erscheint in unregelmäßiger Reihenfolge.

Jede Nummer erscheint in digitaler Version unter  
<https://www.wiso.uni-hamburg.de/de/forschung/working-paper-series/>

### Kontakt:

WiSo-Forschungslabor  
Von-Melle-Park 5  
20146 Hamburg  
E-Mail: [experiments@wiso.uni-hamburg.de](mailto:experiments@wiso.uni-hamburg.de)  
Web: <http://www.wiso.uni-hamburg.de/forschung/forschungslabor/home/>



# Testing the Endowment Effect for Default Rules\*

Isabel Marcin<sup>§</sup> and Andreas Nicklisch<sup>‡</sup>

December 20, 2013

## Abstract

This paper explores potential endowment effects of contractual default rules. For this purpose, we analyze the Hadley liability default clause in a model of bilateral bargaining of lotteries against safe options. The liability default clause determines the right for the safe payoff option. We test the model in series of laboratory experiments. The results reveal a substantial willingness-to-accept to willingness-to-pay gap for the right to change lotteries against safe options. Even if we apply the incentive compatible Becker-DeGroot-Marschak value elicitation mechanism, there is a significant gap indicating a robust endowment effect caused by default rules. Differences of expected values of the lotteries and the safe options consistently decrease the gaps. Implications for applications of default rules in the law are discussed.

Keywords: *Default rules, Endowment effect, Lotteries*

JEL-Classification: C91; K00; K12.

---

\*We gratefully acknowledge the many useful comments provided by Stefan Bechtold, Christoph Engel, Oliver Kirchkamp, Matthew Rabin, and Alessandro Lizzeri. We also thank participants of the IMPRS Ph.D. seminar, the Cleen Conference, Bergen, Norway, and the ECORE Summer School, Leuven, Belgium, for helpful comments. We are indebted to the University of Hamburg for financial support.

<sup>§</sup>*Corresponding author:* Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str.10, 53113 Bonn, Germany, Email: marcin@coll.mpg.de

<sup>‡</sup>University of Hamburg, School of Business, Economics and Social Science, von-Melle-Park 5, 20146 Hamburg, Germany, & Max Planck Institute for Research on Collective Goods, Email: andreas.nicklisch@wiso.uni-hamburg.de

# 1 Introduction

Default rules represent an important element of (contract) law. They determine contractual obligations as long as the contractual parties do not explicitly override them by any other agreement. As default rules are non-mandatory, their importance for contracts appears rather weak. However, it has been shown in the literature that defaults significantly influence behavior in a wide range of fields such as savings plans (Madrian and Shea, 2001; Carroll et al., 2009; Choi et al., 2004), choice of insurance policy (Johnson et al., 1993), or organ donation (Johnson and Goldstein, 2003). Therefore, the interplay between default clauses and economic consequences has been discussed intensively by many scholars of law and economics (Korobkin, 1998b,a; Arlen et al., 2002; Sunstein, 2002; Sloof et al., 2007).

If default rules do not simply distribute entitlements which can be circumvented by contract, but influence crucially stated preferences, the innocence of default rules is put into question. We therefore experimentally analyze one potential problem caused by default clause, the endowment effect. The endowment effect describes the observed divergence between the so-called maximal “willingness to pay” (WTP) to obtain an entitlement and the minimal amount one is willing to accept for a commensurate loss, “willingness to accept” (WTA) (Horowitz and McConnell, 2002; Knetsch et al., 2001). In other words, the owner of a (legal) entitlement values the entitlement more than the non-owner. The (economic) consequences of such biased valuations are far-reaching. When default clauses evoke endowment effects, the final allocations of contracting and trading entitlements need not be efficient (even in the absence of transaction costs). Rather, initial distributions of rights and obligations according to default clauses are likely to be sticky (Schwab, 1988; Korobkin, 1998b; Arlen et al., 2002; Sloof et al., 2007).

While the endowment effect, as measured as a  $WTA/WTP$  gap, is one of the most acknowledged phenomena of behavioral economics (Camerer et al., 2003, p. 15; Plott and Zeiler, 2005, p. 531), economic research still provides inconclusive results. Some scholars suggest that the effect is context-dependent because the commodity used in the experiment (e.g., mugs, lotteries or money) affects the magnitude of the effect (Shogren et al., 1994). Others propose that the effect diminishes with market experience (List, 2003), while more recent literature argues that the magnitude of the gap also depends on experimental procedures (List and Gallet, 2001; Plott and Zeiler, 2005, 2007, 2011; Sugden et al., 2013). Plott and Zeiler (2005) suggest incorporating a list of experimental procedures into WTA-WTP experiments, which reduce likely misconceptions.

This line of research supports the importance of explicitly testing contrac-

tual default rules in a controlled experiment. Taking up the above-mentioned critique, we model the good, contractual default rule, in an abstract setting, yet incorporating three main features. First, a legal rule is an intangible asset. Therefore subjects receive a written description of the rule in place. In contrast to an ordinary good, such as a mug, they have no tangible asset in their hands. Second, the form of ownership attributed by a legal rule depends on the content of each rule. Therefore, we consider a specific rule in our analysis, the Hadley liability default rule. The Hadley rule originates from the case *Hadley v. Baxendale*<sup>1</sup>: A miller contracted a carrier to deliver a broken crank shaft to his mill, which was required as a model to build a new shaft. As the shipment was delayed, the miller sued the carrier for the subsequent lost profits while the mill was inoperative Korobkin (1998b). The court’s decision held in this case in favor of the carrier, establishing the common-law precedent of a default rule which guarantees liability only for foreseeable consequential damages.<sup>2</sup> Third, default rules differ in the value they assign to the contracting parties. We introduce value differences on two dimensions. On the first dimension, we apply default contracts of different expected value. On the second dimension, we vary the difference in expected value of the default and the alternative contract. This allows us to reproduce a variety of contracts.

To avoid other potential effects or misconceptions, we give subjects full information on the expected value of the liability rules, although in real-world situations, the value of a legal contract is often subject to uncertainty. Thereby we can rule out any ambiguity effects. In addition, in our experiment there is no legal uncertainty on the foreseeability of damage and we only consider “foreseeable” damages. According to the Hadley doctrine, the carrier is liable for “foreseeable” damages, but not for “unforeseeable” damages. The Hadley rule specifies that liability for unforeseeable damages must be explicitly included in the contract in advance. Nevertheless, in reality contracting parties may assess differently ex post which damage was foreseeable. That is why eventually a judge needs to decide on that matter. This additional complexity will be waived in favor of keeping the experimental design simple. Consequently, our study analyzes default rules under perfect information. We further exclude any negotiations that would naturally occur when contracting parties conclude a contract, as negotiation may induce

---

<sup>1</sup>156 Eng. Rep. 145 (1854).

<sup>2</sup>The court decided against the miller: only the foreseeable consequential damages at the time of contracting should be awarded to the miller. Yet, the court decided that the consequential damage was not foreseeable for the carrier, as he had not been informed about the use of the broken crank shaft (Ayres and Gertner, 1989, p. 101, Korobkin, 1998b, p. 616).

further strategic behavior.

The default rule in our experiment is modeled in the following way. There is a contract between two parties. One party, the sender, delivers a good to the other party, the receiver. The delivery may be delayed according to an exogenous and commonly known probability. Two different liability rules are introduced. The first rule constitutes a full liability rule: the sender has liability for the delayed delivery and covers the receiver's complete costs for damages. The second rule, motivated by the Hadley limited liability rule, represents the reverse case: the sender has no liability for the delayed delivery and the receiver covers the costs for damages. One of the two liability rules is given as default rule to them.

Consequently, the decision between alternative liability rules is modeled as lotteries against a safe option. Under the limited liability rule, the receiver has a lottery combining a small payoff (delayed delivery) and a large payoff (in-time delivery), while under full liability the large payoff for the receiver is guaranteed for sure. In the experiment, we elicit the valuation for a change from the full to the limited liability rule and vice versa. As such, we follow the tradition of endowment effect experiments using safe payoffs in comparison to lotteries as the trading commodity (e.g., Eisenberger and Weber, 1995). Thus, we can compare our results to other studies on the endowment effects of lotteries.

In the first treatment (hereafter denoted as baseline), we model an intuitive setting to reveal the subjects' valuation for the default rule. In the first case, we elicit the sender's willingness to pay for a liability reduction and the receiver's willingness to accept this reduction, while we elicit, in the second case, the sender's willingness to accept a liability extension and the receiver's willingness to pay for this extension. If the WTP exceeds the WTA, the liability changes accordingly. Comparison between the WTA and the WTP of senders (receivers) under both rules reveals potential endowment effects caused by default rules. However, as we will show below, there is a strategic incentive in the first case (the second case) for the receiver (the sender) to state a WTA which is higher than his/her true evaluation.

Therefore, we compare our results with data from a second experimental setting (hereafter denoted as variation) applying the Becker-DeGroot-Marschak value elicitation mechanism (BDM). This mechanism eliminates strategic incentives for WTA statements, however, at the cost of a highly abstract design. Our twofold design aims at measuring potential endowment effects of default clauses in two "extreme" experimental environments: on the one hand, our baseline treatment allows us to identify biased valuations of senders and receivers in a realistic scenario which may be strategically biased. On the other hand, the scenario in the variation treatment is much

more artificial, but assures the sincere elicitation of stated preferences. Unlike previous research on hypothetical choices under the Hadley rule by law and economic scholars (Korobkin, 1998b), we provide financial incentives, sufficient practice and anonymity, and – at least in variation – an incentive-compatible mechanism.

The most obvious finding to emerge from this study is that we observed significant *WTA/WTP* gaps for default rules consistent over time. Neither the experience accumulated over time nor the value elicitation method appear to influence the gap significantly. Median gaps of 1.31 and 1.45 (baseline and variation treatment) are similar to results found for endowment experiments on lotteries.<sup>3</sup> The difference in expected value of the contracts was found to affect the gap negatively. We interpret this finding in such a way that higher differences in expected value are more salient to subjects so that they elaborate more consciously on their valuations and are, in turn, less affected by the endowment effect. By making use of the isolating power of experiments, other factors that may affect valuations such as strategic effects or the status quo bias can be excluded by the experimental design.

Of course, limitations apply to our study. For instance, we only consider situations where the expected payoff and the probabilities of damage are commonly known. Quite often, this is not the case in reality.<sup>4</sup> Nonetheless, we interpret our experimental analysis and results as a conservative estimate for the isolated endowment effect of default clauses for senders' and receivers' preferences.

The remainder of this paper proceeds as follows. Section 2 embeds our study into the existing literature on the endowment effect and reviews previous empirical research on the conjuncture of default rules and endowment effect. In Section 3, the design of the experiment is explained and subsequently theoretical and behavioral predictions are presented. Section 4 presents and discusses the experimental results of our experiment. Section 5 concludes the paper.

---

<sup>3</sup>E.g., Eisenberger and Weber, 1995), even when trading lotteries with unknown probabilities (i.e., ambiguous commodities; see, e.g., Trautmann and Schmidt, 2012) or for trades between risky commodities (i.e., players stated their *WTA* and *WTP* for not fully reducing the risk of the lottery; see, e.g., Schmidt and Traub, 2009).

<sup>4</sup>Notice that Eisenberger and Weber (1995) find no statistically different *WTA-WTP* ratios for risky and uncertain lotteries, so that a separate analysis of risky environments and uncertain environments appears unnecessary.



## 2 Related Literature

The literature names mainly three functions for default rules. First, they fill gaps of incomplete contracts since contracts may not include terms for all possible contingencies. In case a contingency arises that is not settled in the contract, a default rule will be applied. Second, default rules lower transaction costs between the contracting parties and relieve the negotiation process by providing standard rules. Third, they provide guidance for the design of contractual arrangements (Bechtold, 2010, p. 14). Assuming zero transaction costs, none of these functions conflict with efficient trading. As utility maximizers, contracting parties should agree to those contract terms that maximize their joint utility regardless of any default rule.

With positive transaction costs, the efficiency of the default rule hinges on its content since it affects the probability that the default rule will be kept by the contracting partners.<sup>5</sup> However, even if default rules save transaction costs, this comes at the cost of potential distortion of stated preferences. Two main lines of argument can be identified that explain how default rules may affect valuations. The first one puts forward that strategic considerations may prevent agents from truthfully revealing their true preferences. Some reasons for this behavior may include strategic effects that arise due to relational factors and norms (Bernstein, 1993), the fear of adverse inferences (Spier, 1992), strategic behavior (Ayres and Gertner, 1989), or externalities (Kahan and Klausner, 1997).

The second stream claims that psychological factors can alter individual valuations (Korobkin, 1998b; Sunstein, 2002). The status quo bias, for example, shows that people systematically favor maintaining a state of affairs that they perceive as being the status quo, instead of switching to an alternative state (Kahneman et al., 1991). The endowment effect also belongs to the group of psychological factors that shifts preferences in favor of the default rule. According to Korobkin (1998b) and Sunstein (2002), an endowment effect occurs if agents perceive the rights guaranteed by the default rule as their entitlement. This can evoke loss aversion within the agents.<sup>6</sup> Based on

---

<sup>5</sup>Two approaches for the design of default rules have been discussed extensively in the literature. According to the majoritarian approach (also called the “what parties would have wanted” default), default rules shall represent the majority’s interest to minimize aggregate transaction costs. In contrast, the minoritarian approach does not assume transaction costs to be uniformly distributed so that a minoritarian-based default rule could also minimize aggregate costs. For more information, see, for instance, Ayres and Gertner (1989); Geis (2005); Bechtold (2010).

<sup>6</sup>Alternatively, scholars such as Arlen and Tontrup (2013) argue in favor of a process account of the endowment effect. The endowment effect is driven by the fear of regretting one’s transaction. The anticipated regret is higher the more one feels responsible.



this aversion, the endowment effect predicts that the agents' *WTA* to forgo the default rule will be higher than their *WTP* for it. Our experiment specifically addresses this effect while excluding other psychological or strategic factors.

The endowment effect is non-universal in the sense that it varies considerably in size between individuals, depending the commodity traded and the experimental procedures.<sup>7</sup> Two different main elicitation methods to measure the *WTA/WTP* gap and a large variety of experimental variations can be found in the literature. The first method elicits valuations for particular goods by asking for *WTA* and *WTP* in contingent valuation studies which mostly involve hypothetical choices (for instance Kahneman et al. 1986). Most often these were open-ended questions with hypothetical choices (Horowitz and McConnell, 2002, p. 426). These studies are criticized because subjects face no real incentives since they cannot earn money. Also, they may strategically overstate their *WTA* and understate their *WTP* (Posavac, 1998) or fall prey to common bargaining habits like selling high and buying low (Horowitz and McConnell, 2002; Sayman and Öncüler, 2005; Plott and Zeiler, 2005).

The second method uses laboratory experiments where valuations are elicited with mostly incentive-compatible mechanisms, such as the BDM, Vickrey second-price or random “n-th” price auction (List and Gallet, 2001).<sup>8</sup> While incentive-compatible mechanisms are preferable from a theoretical viewpoint, empirical results are mixed. Horowitz and McConnell (2002) find evidence in their meta-study that incentive-compatible mechanisms yield higher *WTA/WTP* gaps than other methods, Sayman and Öncüler (2005, p. 303-305) detect lower *WTA/WTP* gaps in their meta study.

Past research has also shown the exigency to control for experimental variations. Plott and Zeiler (2005, 2007, 2011) claim that the gap reflects subject misconceptions which are evoked through specific experimental procedures as well as through variations of classical preference theory (such as signaling effect,<sup>9</sup> other-regarding preferences, or cascade theories). In two experiments, Plott and Zeiler (2005, 2007) find no significant *WTA/WTP* gap when they apply an incentive-compatible mechanism, training, prac-

---

<sup>7</sup>Numerous studies found different *WTA/WTP* gap sizes. For more information, see a meta study on the endowment effect by Horowitz and McConnell (2002).

<sup>8</sup>There is also a third mechanism that allows for testing the endowment effect. In so-called exchange experiments, valuations are not directly elicited, but the volume of trade is measured. See, for example, Knetsch (1989), who found a very low volume of trade.

<sup>9</sup>Signaling theories suggest that individuals may interpret the initial entitlement that was chosen by the experimenter as a signal of relative quality (Plott and Zeiler, 2007, p. 1450).

tice/coaching, and anonymity.<sup>10</sup> As a consequence, one could conclude that gaps found in other studies that lack these devices may be based upon misconceptions and not upon psychological characteristics such as loss aversion. Plott and Zeiler’s design aroused approval (Bardsley et al., 2009, p. 181-182) as well as rejection (Isoni et al., 2011; Knetsch and Wong, 2009)<sup>11</sup> and their procedures have been applied by other studies, which show mixed results.<sup>12</sup> For our experimental test, we will take the important aspects of Plott and Zeiler’s design into consideration.

Parts of the legal literature, particularly those at the intersection of law and economics, have already conducted experiments on the endowment effect of default rights. One of the earliest experiments by Schwab (1988) measured the valuations for a legal right, the relocation of work to nonunion plants, and found no significant gap.<sup>13</sup> A more recent experiment by Sloof et al. (2007) addresses default breach remedies in an abstract game. In a bilateral bargaining game, they find a significant effect of default rules on outcomes. However, the reason is not that proposals were biased towards the default, but whenever parties disagreed about the preferred contract, the default rule was applied by experimental design. This led to a prevalence of the default contract.

Sunstein (2002) also tested a default rule in the context of employment contract negotiations. He asked law students to imagine that they are negotiating their contract to work with a law firm. The students were divided into two groups where each of them received different information on the default holiday length. He elicited the value of holiday by asking for the corresponding salary reduction/improvement to extend/shorten holiday length and found a considerate WTA-WTP gap.<sup>14</sup>

One of the best-known legal experiments is Korobkin’s (1998b) experiment on default rules, which also serves as a blueprint for our experiment. With the help of three controlled experiments, Korobkin (1998b) tests the

---

<sup>10</sup>They also quote paid practice as a procedure, but find in their experiment that payments in the practice rounds are not necessary.

<sup>11</sup>Knetsch and Wong (2009) argue that Plott and Zeiler’s procedures not only nullified the effect of subjects’s misconceptions, but also changed the reference state. For instance, by only showing the subjects the good instead of giving them the good, subjects might not have felt endowed with the good. Thus, Plott and Zeiler could not measure an “endowment” effect.

<sup>12</sup>Consistent results were found by Kovalchik et al. (2005), while non-consistent results were detected by Isoni et al. (2011) and Knetsch and Wong (2009).

<sup>13</sup>Other scholars discuss on the basis of theoretical models if this Hadley rule is efficient (Bebchuk and Shavell, 1991; Adler, 1999).

<sup>14</sup>Further experiments on legal default rules include Arlen et al. (2002) and Buccafusco and Sprigman (2011).

effect of the liability rule. Subjects participate in a survey taking the hypothetical role of a lawyer providing advice to a client on the negotiation of a commercial contract (the example is designed similarly to the Hadley v. Baxendale case). Subjects are asked to evaluate a contract design with a choice of two liability clauses. Option one insures against all consequential damages and is thus in favor of the client, while option two is in favor of the opposing party. Students are told which approximate cost would accrue under the two possible liability clauses to the client. Subjects are then divided into two groups. One group is told that option one is statutory law, while the other group is told the equivalent for option clause two. When eliciting their valuations for a change to the other liability rule, Korobkin finds that both groups favor their default rule. He attributes this result to loss aversion and argues that subjects fear ex-post regret from actively changing the contract terms. This experimental evidence is important, as transaction costs can be excluded as an obstacle to negotiations. Thus, the observed *WTA/WTP* gap is not driven by transaction costs. However, in our view, some limitations apply to this experiment. It only elicits hypothetical choices which may be subject to strategically over- and underbidding, and a no incentive-compatible mechanism is applied.

From our review, we would like to formulate two crucial demands when testing the endowment effect in a legal context. On the one hand, there is still a need to conduct more real choice experiments with incentive-compatible mechanisms. On the other hand, the review underlines the importance of using experimental controls. Consequently, our experiment tests the endowment effect for the Hadley liability rule alongside the idea of Korobkin’s hypothetical choice experiment. In addition, it integrates experimental controls (such as use of an incentive-compatible mechanism, practice rounds, repeated rounds for experience). In contrast to Korobkin’s experiment, subjects will not only assume a hypothetical advising role, but they will participate in a real choice experiment with a performance-oriented payment. Risk with regards to the value of the liability rule is also integrated (by allowing damage to occur with a certain probability). Finally, our experimental design allows testing for a different degree of realism by using two different types of value elicitation methods. The details of the experimental design will be described in the next chapter.

### 3 Experimental Design

To examine whether default rules for contracts evoke an endowment effect, we conducted an experiment with two treatments. Following Korobkin (1998b),

subjects were placed in a framed situation that is similar to the Hadley v. Baxley case. This design is not tailored to a specific application in law, but serves as a framework incorporating important features of many relevant situations.

In the experiment, two different liability rules regulate liability in case of damages. The “liability rule” endows the company with a full cover of potential damages, while the other reflects a “no liability rule”. Treatment conditions vary the elicitation mechanism. The first treatment (*baseline*) studies whether subjects exhibit a *WTA/WTP* gap with reference to the default liability rule in a bilateral bargaining of lotteries. In a second treatment (*variation*), the experiment is embedded into a more abstract setting. This treatment examines if the *WTA-WTP* disparity size changes when subjects play in a unilateral setting using the incentive-compatible BDM procedure. In addition, variations within treatments explore whether differences in expected values of the liability rules affect the *WTA-WTP* disparity, and, as a robustness check, irrational *WTA*- or *WTP*-proposals.

### 3.1 The Game

Let us first consider *baseline*. There are two types of players, 1 and 2. Each player receives some money  $E$ . Players are then assigned to two different regimes: under regime *liability*, player 1 is endowed with contract  $A$ , player 2 with contract  $B$ ; under regime *no liability*, it is the other way around; player 1 is endowed with contract  $B$  and player 2 with contract  $A$ . The contracts are simple lotteries; contract  $A$  provides a degenerate lottery  $(p, X_1; 1-p, X_2)$  with payoffs  $X_1 = X_2 := X$  and probability  $p = 0.5$ ; contract  $B$  is a true lottery  $(p, Y; 1-p, y)$  with payoffs  $Y > y$  and probability  $p = 1-p = 0.5$ .

Before nature determines the lotteries, and subjects are paid out, player 1 states under *liability* her *WTA* with  $WTA \in [0, E]$  to change from contract  $A$  to contract  $B$ , and player 2 states her *WTP* with  $WTP \in [0, E]$  to change from contract  $B$  to contract  $A$  ((under *no liability*) player 1 states her *WTP* to change from contract  $B$  to contract  $A$ , and player 2 states her *WTA* to change from contract  $A$  to contract  $B$ ). *WTA* and *WTP* are compared pairwise. If  $WTA \leq WTP$ , the contracts will be exchanged and player 1 (player 2) receives *WTA*, while player 2 (player 1) pays *WTA* as the price for the exchange. If  $WTA > WTP$  players keep their initial contracts, and no price is paid. Under *liability*, the expected payoffs of player 1,  $\pi_1$ , and player

2,  $\pi_2$ , are

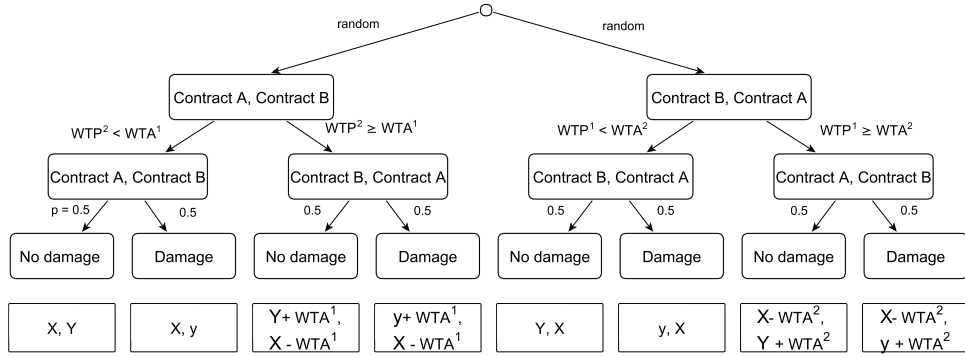
$$\pi_1 = \begin{cases} .5Y + .5y + E + WTA & \text{if } WTA \leq WTP \\ X + E & \text{if } WTA > WTP \end{cases} \quad (1)$$

and

$$\pi_2 = \begin{cases} X + E - WTA & \text{if } WTA \leq WTP \\ .5Y + .5y + E & \text{if } WTA > WTP \end{cases} \quad (2)$$

Under *no liability*, player 1 receives an expected payoff according to equation 2 and player 2 according to equation 1. Figure 1 illustrates the decision tree for *baseline* (except the endowment  $E$ ).

Figure 1: Decision tree for *baseline*



Player 1's payoffs are reported first, player 2's payoffs second;  $WTA^1$  ( $WTP^2$ ) denotes player 1's willingness to accept (player 2's willingness to pay) and vice versa.

As mentioned earlier, the elicitation device used in *baseline* is not fully incentive-compatible, as the player who determines  $WTA$  can increase the payoff by announcing a slightly greater  $\widehat{WTA}$ <sup>15</sup> than her true valuation,  $WTA$ , for contract A. Particularly, player 1 under *liability* and player 2 under *no liability* face the same problem as a bidder in a first-price procurement auction. Her expected surplus from exchanging contract A against contract B,  $\Pi_{WTA}$ , is given by

$$\Pi_{WTA} = (\widehat{WTA} - WTA)F(\widehat{WTP} \geq \widehat{WTA}) \quad (3)$$

<sup>15</sup>In this section,  $\widehat{WTA}$  indicates announced  $WTA$ , while  $WTA$  denotes the true valuation of the player. In the other sections, this distinction will not be applied for simplicity.

with  $F(\cdot)$  denoting the probability that  $\widehat{WTP} \geq \widehat{WTA}$ . Equation 3 reflects the gains the player can make in case the stated  $WTP$ ,  $\widehat{WTP}$ , is larger than or equal to the stated  $WTA$ ,  $\widehat{WTA}$ . She gains the difference between her stated  $\widehat{WTA}$  and her true valuation  $WTA$ . To maximize the expected surplus, the first-order condition of  $\Pi_{WTA}$  with regard to  $\widehat{WTA}$ , denoting  $f(\widehat{WTP} > \widehat{WTA}) := \frac{\partial F(\widehat{WTP} \geq \widehat{WTA})}{\partial \widehat{WTA}}$ , yields

$$\widehat{WTA} = WTA - \frac{F(\widehat{WTP} \geq \widehat{WTA})}{f(\widehat{WTP} \geq \widehat{WTA})} \quad (4)$$

Since the partial differential of the distribution function  $F$  with respect to  $\widehat{WTA}$  is negative,  $f(\widehat{WTP} \geq \widehat{WTA}) < 0$ , one can write

$$\widehat{WTA} = WTA + \frac{F(\widehat{WTP} \geq \widehat{WTA})}{\left| f(\widehat{WTP} \geq \widehat{WTA}) \right|} \quad (5)$$

According to equation 5,  $\widehat{WTA}$  is greater than or equal to  $WTA$ . Hence, in *baseline*, players face the incentive to indicate a  $\widehat{WTA}$  that is larger than their actual valuation.<sup>16</sup>

However, the upward-biased stated  $WTA$  does not affect the other subject's stated  $WTP$ . While the stated  $WTA$  corresponds with a bid in a first-price procurement auction, the  $WTP$  position reflects a bid in a second-price auction: second-price auctions eliminate strategic bidding and induce a truthful revelation of one's valuation of the object (Milgrom, 2004, pp. 9-10). This becomes obvious, when we consider the player's expected surplus from exchanging contract  $B$  against contract  $A$ ,  $\Pi_{WTP}$ :

$$\Pi_{WTP} = (WTP - \widehat{WTA})F(\widehat{WTP} \geq \widehat{WTA}) \quad (6)$$

She can only profit from buying contract  $A$  if the price  $\widehat{WTA}$  is smaller than her actual valuation  $WTP$ . Consequently, for any  $\widehat{WTP} < WTP$  increasing  $\widehat{WTP}$  increases the chance for changing the contracts without increasing its price. On the other hand, stating any  $\widehat{WTP} > WTP$  in comparison to  $\widehat{WTP} = WTP$  increases only the chance for a loss due to  $\widehat{WTP} > \widehat{WTA} > WTP$ . Thus it is the player's dominant strategy to announce her real value. Thus, in *baseline* only the  $WTA$  bid might be slightly

---

<sup>16</sup>Accurate information on the density function  $f(\cdot)$  would allow to define  $\widehat{WTA}$  explicitly.

upward biased; the *WTP* bid should not differ from the actual WTP valuation.

Let us now proceed with the *variation* treatment. The setting of *variation* equals *baseline*'s setting, except that this treatment applies the standard BDM elicitation method for lottery valuations (Plott and Zeiler, 2011; Isoni et al., 2011). That is, the BDM mechanism (Becker, 1964) pits each seller and buyer against a random device. The BDM mechanism is incentive-compatible revealing the true value of individual WTA and WTP values. Players endowed with contract *A* state their *WTA*. If the stated *WTA* does not exceed a random offer *RO*, the player receives contract *B* in exchange for contract *A* and the price *RO*. Players endowed with contract *B* state their *WTP*. If the stated *WTP* exceeds or equals the random offer *RO*, the player receives contract *A* in exchange for contract *B* and pays the price *RO*.<sup>17</sup> Notice that random offers are uniformly distributed on a commonly known range, so that each price occurs with equal probability. Therefore, under *liability*, the payoffs of player 1,  $\pi_1$ , and player 2,  $\pi_2$ , are

$$\pi_1 = \begin{cases} .5Y + .5y + E + RO & \text{if } WTA \leq RO \\ X + E & \text{if } WTA > RO \end{cases} \quad (7)$$

and

$$\pi_2 = \begin{cases} X + E - RO & \text{if } WTP \geq RO \\ .5Y + .5y + E & \text{if } WTP < RO \end{cases} \quad (8)$$

Under *no liability*, player 1 receives a payoff according to equation 8 and player 2 according to equation 7. Under the assumption of stable preferences, constant absolute risk aversion,<sup>18</sup> and a utility maximization strategy,

---

<sup>17</sup>Some scholars have raised theoretical objections about the incentive-compatibility of the BDM procedure. For example, Karni and Safra (1987) have shown that the dominant strategy of revealing one's true value fails to hold for preferences that cannot be expressed by expected utility. Noussair et al. (2004) compare the BDM mechanism and another incentive compatible mechanism, the Vickrey auction. They find that the Vickrey auction is more effective as an (induced) value elicitation device than the BDM process. However, the BDM mechanism remains widely used.

<sup>18</sup> According to Isoni et al. (2011, p. 997), in expected utility theory it is common use to assume that absolute risk aversion decreases with increasing income. However, for experimental purposes an approximation through constant absolute risk aversion is often made because differences in wealth are small and the expected change in risk aversion is negligible. In addition, there is experimental evidence that the WTA-WTP gap still remains significant even if a reversed income effect is present (e.g., Schmidt and Traub, 2009).



according to game-theoretical predictions, the subjects' WTP for the alternative liability rule should be equal to their WTA.

### 3.2 Experimental Procedure

Our experiment is explicitly framed in the legal context: a company (player 1) and a deliverer (player 2) form a contract on the shipping of a crucial good. The delivery incurs the risk of a delay (with a constant probability of 50%) and resulting damages for the company. The game is played repeatedly over 17 rounds, with the value of the good and the damage varying in each round.<sup>19</sup> In each round, either the regime *liability* or the regime *no liability* is effective. That is, either player 1 or player 2 is endowed with the safe option (i.e., the degenerated lottery). Then, players may exchange lotteries.

The 17 rounds of the experiment are divided into three, commonly known stages: a practice stage (three practice rounds), a low-stake stage (six rounds), and a high-stake stage (eight rounds). The 14 non-practice rounds contain seven different contract pairs. Each contract pair has the same value for the good and for the damage (i.e., the expected value of lotteries and spread of payoffs in lottery according to contract  $B$ ). However, subjects are assigned the opposite default regimes in the two rounds. Therefore, we have for each player both the  $WTA$  and the  $WTP$ , allowing us to test for within-player  $WTA/WTP$  gaps.

Table 1 shows all seven contracts pairs of the experiment (excluding practice), the differences of expected values between contract  $A$  and contract  $B$  ( $E(A) - E(B)$ ), and the maximal rational offer ( $Max$ ). In the following, all variables will be explained.

As illustrated in Table 1, there are three low-stake contract pairs for which  $WTA$  and  $WTP$  valuations are elicited in the first six rounds, and four high-stake contract pairs for which  $WTA$  and  $WTP$  valuations are elicited in the subsequent eight rounds. In the low-stake stage subjects receive an endowment of  $E = 0.5$  experimental currency unit (ECU), while they receive 5 ECU in the high-stake rounds. To avoid potential order effects, the order of the rounds within the low-stake and high-stake stages are drawn at random.

Each contract pair is played twice in order to allow for within-subject comparisons of  $WTP$  and  $WTA$ . For instance, once the subjects are endowed with contract  $A$  and another time with contract  $B$ . The next-to-last column

---

<sup>19</sup>Payoffs of contracts differ only slightly and probabilities are kept at 0.5 to reduce possible misconceptions. Plott and Zeiler (2005, p. 540) argue that misconceptions are driven by a lacking understanding of probability. Therefore, only lotteries with equiprobability are used as we assume that subjects feel more at ease comparing two potential lottery outcomes of equiprobability.

Table 1: Lottery design

	Contract pair	Lottery <i>A</i>		Lottery <i>B</i>		$E(A)-E(B)$	<i>Max</i>
Low-stake $E=0.5$	1	1	1	1	0.5	0.25	0.5
	2	1.5	1.5	2	1	0	0.5
	3	1	1	1	0.6	0.2	0.4
High-stake $E=5$	4	10	10	10	5	2.5	5
	5	15	15	15	10	2.5	5
	6	10	10	12	8	0	2
	7	5	5	7	3	0	2

of table 1 shows the difference between the expected values of contract *A* and *B*,  $E(A)-E(B)$ . Some contract pairs (1, 3, 4, 5) have a positive difference between expected values. In these contract pairs, contracts *A* and *B* differ not only with regard to the risk involved in the lottery, but contract *A* leads to an expected payoff increase. Those lotteries map exactly the Hadley liability default rule: liability guarantees a compensation of damages, that is, contract *A* guarantees the larger payoff of the two possible outcomes of contract *B*. Since there is an ongoing discussion in the literature whether offering both less risk and more payoffs systematically influences the size of the endowment effect (Schmidt and Traub, 2009), we also test contract pairs with no differences in expected payoffs. Therefore, we can interpret the *WTP* of contract pairs 2, 6, and 7 (of course, also *WTA* in *variation*) as certainty equivalents and will thus be a measure for risk aversion. We will come back to this point below.

The last column of table 1 provides the upper bound of rational offers (*Max*) representing the maximal rational offer. The lower bound for all rounds is zero. The maximal rational offer represents the value that fully risk-averse subjects would just be willing to pay. It is defined by the maximal difference between the payoff of the degenerate lottery in contract *A* and the payoffs of lottery in contract *B*.<sup>20</sup> Some contract pairs (2, 6, and 7) therefore allow conducting rationality checks. In these rounds, the maximal rational offer is lower than the endowment, so that subjects can make a higher than rational offer. Not all rounds allow for rationality checks, because we try to design a range as narrowly as possible to limit the risk the subjects face concerning the final price they will eventually pay.

<sup>20</sup>For instance, if the degenerate lottery in contract *A* provides a payoff of  $X$ , while the lottery in contract *B* has a 50 percent probability of paying  $Y$  and a 50 percent probability of  $y$  with  $Y > y$ , the maximal rational offer is defined by the absolute value of  $\max(X - Y, X - y)$ . Valuations above this value violate rationality.

In the following, the general procedure of each treatment is explained. The experiment was conducted in the laboratory for economics and social science at the University of Hamburg in September 2012. It was programmed and conducted with the software z-Tree (Fischbacher, 2007). The experiment was organized and recruited with the software hroot (Bock et al., 2012). A total of 107 students from various disciplines participated in the experiment.

Eight sessions, divided into five *baseline* sessions and two *variation* sessions, were run. Prior to round one, subjects were randomly allocated their role; either they were player 1 (company) or player 2 (deliverer). Roles remained constant throughout the entire experiment. 90 subjects participated in *baseline* and 17 subjects participated in *variation*. In the *baseline* treatment, subjects were randomly matched in pairs from matching groups of six subjects. This alternating allocation prevented subjects from learning too much about the other subjects' behavior. In our analysis, each group of six players represents one independent observation, leaving us with 15 independent observations in *baseline*. In the *variation* treatment, every individual represents one independent observation, yielding 17 independent observations in *variation*.

At the beginning of each experimental session, subjects enter the laboratory where each subjects draws a number card that determines their booth. The random drawing ensures anonymity. Subjects are guaranteed that no other subjects will be able to identify their actions or payoffs. Subsequently, written general instructions (see Appendix B) are distributed to the subjects and read out loud by the experimenter. The legal context is introduced to the subjects by the written instructions, which frames the game as a shipping contract under liability or no liability. The general instructions will present the subjects with the game and will give numerical examples to provide subjects with concrete illustrations of the WTA-WTP elicitation procedure.

To account for misconceptions, as explained in Plott and Zeiler (2005, p. 540), we include various simplifications and controls. To familiarize the subjects with the game and the value elicitation mechanism, we employ instructions with examples, control questions which need to be answered correctly, and three training rounds.

Upon arrival, subjects received a non-performance-related show-up fee of 5 Euros, and the amounts earned from the experiment were exchanged at a rate of 4 ECU = 1 Euro. The sum of the show-up fee and the amounts earned from two randomly determined rounds (one low-stake and one high-stake round) determined the final payoff (to avoid effects of accumulated profits) which averaged 9.10 Euros and ranged from 7.60 Euros to 10.50 Euros. On average, one session lasted 60 minutes.

### 3.3 Game-theoretical and Behavioral Predictions

The following section on behavioral predictions shall provide an overview of expected subject behavior and expected changes in behavior between *baseline* and *variation*. Of course, game theory predicts that a subject's WTP for the alternative liability rule is different from the WTA in *baseline*. Indeed, experimental data relying on a similar design as our *baseline* – but without monetary incentives for participants – by Korobkin (1998b) show a significant *WTA/WTP* gap. The interesting question is whether there we find an endowment effect for the default clause in *variation*. In *baseline*, the endowment could amplify the *WTA/WTP* gap resulting from the incentives of the elicitation mechanism; in *variation*, the endowment effect could create the gap. The effect emerges because the subject's valuation of the contract terms changes once the default contract becomes part of the individual's endowment. The individual then attributes a higher value to the entitlement (Thaler, 1980). However, as explained in the literature review, it is not clear if contract default rules will evoke such a change in valuation, as default rules only influence the owner's payoff probabilistically, that is, they materialize only in the bad state of nature yielding the small payoff  $y$ . As it is still debated in which contexts and for which items such a disparity exists, our experiment aims to contribute to the literature by testing for the existence of *WTA/WTP* gaps for default rules.

**Hypothesis 1.** On average, the subjects' *WTA* is higher than the subjects' *WTP* for a contract change in *baseline* and *variation*.

In particular, we want to obtain insights into which factors affect the size of the *WTA/WTP* gap. For this purpose, a further two testable hypotheses will be examined. The first one addresses potential sources of the endowment effect, especially the role of misconceptions. They can be caused by a failure to understand the elicitation mechanism (Plott and Zeiler, 2011, online appendix). In turn, experience might decrease the size of the effect. By playing a game for several rounds, subjects should learn to assess the consequences of their choices and gain confidence in evaluating the contracts.<sup>21</sup> If the *WTA/WTP* gap is indeed caused by misconceptions, it is likely that responses in earlier rounds are more affected than those in later rounds.

By contrast, experience may also increase the *WTA/WTP* gap. Earlier results suggest that subjects frequently evaluate the regret as greater when the unwanted consequences follow from action rather than from inaction and the subjects might therefore fear ex-post regret (Korobkin, 1998a). In our

---

<sup>21</sup>The field experiment by John List (2003) supports this hypothesis.

experiment, unwanted consequences triggering ex-post regret could follow losses induced by a change from the sure contract  $A$  to the risky contract  $B$ . This may result in a higher  $WTA$ . On the other hand, we do not expect a higher  $WTP$  for subjects initially endowed with contract  $B$  as the unwanted consequences that may appear if subjects keep this contract would follow from inaction. In order to control for practice and/or regret effects, the influence of time on the  $WTA/WTP$  gap will be evaluated in a regression analysis testing:

**Hypothesis 2.** A  $WTA$ - $WTP$  disparity exists, but decreases over time.

Finally, we want to analyze whether variations in the payoff structure of the contracts influences the  $WTA/WTP$  gap size. If subjects form their preferences for the contracts on the spot, they may be influenced by cues of the decision problem. Peters et al. (2003) explain a  $WTA/WTP$  gap by the salience of different numeric cues. A large difference in expected value or the change from low- to high-stake rounds appear salient to the subjects. With higher salience, subjects are more likely to make a high effort when elaborating their valuations than subjects in rounds with contracts of the same expected value. This may reduce the influence of a psychological bias, such as the endowment effect. One could therefore assume that the endowment effect is less pronounced in rounds with large differences in expected values, and once subjects decide in high-stake rounds. Of course, a result showing that endowment effects decrease with increasing monetary consequences (i.e., expected payoff differences, high stakes) may be a positive sign for law-making. Subjects' decisions on fundamental problems will be less affected by the endowment effect. Hence, we examine

**Hypothesis 3.** The difference in expected values between contracts and high stakes negatively affects the  $WTA/WTP$  gap size.

## 4 Experimental Results

This section will intensively discuss the results of the experiment using statistical techniques. To provide an overview, the most important results will be briefly summarized. As a main result, a disparity between  $WTA$  and  $WTP$  is found throughout all treatment variations persistent over time, which affirms Hypothesis 1 and rejects Hypothesis 2. The value elicitation method does not significantly alter the gap. However, we can confirm Hypothesis 3, as we find that the difference in expected value between the contracts  $A$  and  $B$  negatively affects the gap. These results are derived based on descriptive as well as inferential statistics.

To support the results above, this chapter proceeds in the following way. The first section presents descriptive statistics and non-parametric tests. In addition, a second subsection on potential caveats explores whether within gaps differ from between gaps, and a robustness check for irrational responses is also performed. The third part uses a regression analysis to estimate the effects of treatment variations. As a conclusion, the section will assess the previously stated hypotheses.

## 4.1 Descriptive Statistics and Non-parametric Tests

In the following, we present the descriptive statistics of *baseline* and *variation*. Table 2 summarizes *WTA* and *WTP* for contract pairs 1 to 7 in *baseline* and shows the number of independent observations ( $n$ ) for each valuation, the difference between the expected value of the contract lotteries ( $\Delta$  EV), the mean, the median and the standard deviation (SD) of participants' reported valuations. As mentioned earlier, we take the mean of each matching group of six players as one independent observation. As there are two types of players in this treatment whose valuations for the same contract pair are elicited in different rounds, one independent observation is constructed by taking the arithmetic mean of the valuations given by the three player 1 subjects elicited in one round, as well as the valuations given by the three player 2 subjects elicited in the corresponding round.<sup>22</sup>

For each contract pair, the table also shows mean and median of the "standardized *WTA/WTP*" disparities. Standardized *WTA/WTP* disparities are calculated as ratios following Harless (1989), Plott and Zeiler (2005), Schmidt and Traub (2009), and Isoni et al. (2011). The results of the Wilcoxon signed-rank test are reported below the *WTA* and *WTP* valuations. It tests under the null hypothesis that the two corresponding median *WTA* and median *WTP* valuations are identical. A non-rejection of the test indicates that *WTA* and *WTP* valuations do not significantly differ from each other.

For the low-stake contract pairs, the mean *WTAs* range from 0.34 to 0.38 ECU (median 0.35 - 0.38), while the mean *WTPs* range from 0.28 to 0.3 ECU (median 0.28 - 0.3). Corresponding results for the high-stake contract pairs

---

<sup>22</sup>For example, in one specific round, say 1, contract pair 3 is randomly chosen to be played. Player 1 subjects are endowed with contract A and indicate their *WTA*, while player 2 subjects are endowed with contract B and state their *WTP* (and vice versa in another round, say 2). Consequently, the matching group *WTA* mean is constructed by taking the average of the *WTA* valuations of player 1 subjects from round 1 and the *WTA* valuations of player 2 subjects from round 2 (and vice versa for the matching group *WTP* mean).

Table 2: Summary of *baseline*

Contract pairs	Low-stake				High-stake		
	1	2	3	4	5	6	7
<i>WTA</i>							
n	15	15	15	15	15	15	15
$\Delta$ EV	0.25	0	0.2	2.5	2.5	0	0
Mean	0.36	0.38	0.34	3.67	3.68	2.48	2.61
Median	0.35	0.38	0.36	3.83	3.78	2.58	2.46
SD	0.05	0.07	0.06	0.51	0.49	0.68	0.60
<i>WTP</i>							
n	15	15	15	15	15	15	15
$\Delta$ EV	0.25	0	0.2	2.5	2.5	0	0
Mean	0.30	0.27	0.28	2.72	2.82	1.77	1.71
Median	0.30	0.28	0.28	2.65	2.71	1.75	1.71
SD	0.05	0.08	0.06	0.49	0.46	0.63	0.42
Significance: $WTA=WTP^1$	***	***	***	***	***	***	***
<i>WTA/WTP</i>							
n	15	15	15	15	15	15	15
Mean	1.20	1.70	1.25	1.37	1.32	1.56	1.58
Median	1.16	1.34	1.22	1.30	1.36	1.48	1.49

<sup>1</sup> Test based on signed ranks (two-tailed Wilcoxon signed rank test).

Note: Significance levels are indicated as \*\*\* (1%), \*\* (5%), and \* (10%).

are a mean *WTA* of 2.48 – 3.68 (median 2.46 - 3.83) and a mean *WTP* of 1.71 – 2.82 (median 1.71 - 2.71). The overall mean *WTA/WTP* for *baseline* amounts to 1.43 ECU (median 1.31). For the low-stake contract pairs, the mean amounts to 1.45 (median 1.34), while the mean of the high-stake contract pairs is 1.46 (median 1.40). As shown in Table 2, the median *WTA* is significantly greater than the median *WTP* at the one-percent level in all comparisons. Thus, the results of the Wilcoxon test support the hypothesis that the two dependent samples (*WTA* and *WTP* for each contract pair) were drawn from nonidentical distributions.

In *variation*, the median over all contract pairs is 1.45 (arithmetic mean 2.25), which is 0.11 points higher than the median of *baseline* (mean is higher by 0.82 points).<sup>23</sup> Thus, it can be hypothesized that the more realistic *base-*

<sup>23</sup>The large difference in the mean *WTA/WTP* gap, yet lower difference in the median between treatments can be explained by the greater degree of aggregation in *baseline*. In



*line* treatment led to a decrease of the gap, in comparison to the more abstract variation treatment. However, the difference in the median gap is not significant.<sup>24</sup>

Table 3: Summary of *variation*

Contract pairs	Low-stake			High-stake			
	1	2	3	4	5	6	7
<i>WTA</i>							
n	17	17	17	17	17	17	17
$\Delta$ EV	0.25	0	0.2	2.5	2.5	0	0
Mean	0.34	0.34	0.33	3.75	3.45	2.35	2.48
Median	0.35	0.45	0.3	3.65	3.5	2	2
SD	0.12	0.19	0.12	1.08	1.55	1.9	1.86
<i>WTP</i>							
n	17	17	17	17	17	17	17
$\Delta$ EV	0.25	0	0.2	2.5	2.5	0	0
Mean	0.28	0.15	0.24	2.42	2.77	1.7	1.6
Median	0.25	0.1	0.25	2.5	2.5	1.35	1
SD	0.14	0.13	0.11	1.16	1.74	1.72	1.69
Significance: WTA=WTP <sup>1</sup>		**	**	***	**	**	**
<i>WTA/WTP</i>							
n	16	13	16	16	15	12	15
Mean	1.73	1.9	1.46	2	1.1	1.31	6.14
Median	1.2	2	1.29	1.55	1.02	1.17	2

<sup>1</sup> Test based on signed ranks (two-tailed Wilcoxon signed rank test).

Note: Significance levels are indicated as \*\*\* (1%), \*\* (5%), and \* (10%).

For the low-stake contract pairs of *variation*, mean *WTAs* range from 0.33 to 0.34 ECU (median 0.3 - 0.4), while mean *WTPs* range from 0.15 to 0.28 ECU (median 0.1 - 0.25). For the high-stake contract pairs, mean *WTAs* amount to 2.35 to 3.75 ECU (median 2 - 3.65) and mean *WTPs* to 1.6 to 2.77 (median 1 - 2.5).

*variation* high individual *WTA/WTP* gaps affect more strongly the overall mean than in *baseline*, where we average over six values. Since the median is less affected by extreme values, the median is more adequate to compare both treatments.

<sup>24</sup>We conducted a median test on the null hypothesis that the *WTA/WTP* gaps of both treatments were drawn from distributions with the same median. On the basis of a continuity corrected chi-squared test statistic the hypothesis cannot be rejected at a significance level of 10 percent for any pair of rounds.

The mean  $WTA/WTP$  amounts to 1.68 (median 1.46) for the low-stake contract pairs and 2.70 (median 1.45) for the high-stake contract pairs.<sup>25</sup> The results of the Wilcoxon test show that the median  $WTA$  significantly exceeds the median  $WTP$  in most comparisons at the one- or five-percent level. However, this does not hold true for contract pair 1. The results in *variation* show a lower degree of significance than the ones detected in *baseline* which is probably due to a decreased variation among the independent observations, since we take, in *baseline*, the average over six players as one independent observation.

Overall, a comparison with other studies summarized in the meta study by (Horowitz and McConnell, 2002) yields that our results lie below the median ratio for lotteries (1.66), as well as below the median ratio for ordinary private goods (2.26) (Horowitz and McConnell, 2002, p. 433).<sup>26</sup> Yet, the explanatory power of this comparison is limited, since the studies employ different elicitation mechanisms and some experiments only elicit hypothetical choices. Contrasting our results with other gaps for lotteries from experiments that employ a similar elicitation procedure (such as the BDM procedure), we find slightly lower numbers than our ratio: Plott and Zeiler (2005) report medians of 1.15-1.16, Schmidt and Traub (2009) 1.06, and Isoni et al. (2011) 1.13-1.33 for gaps between  $WTA$  for selling a lottery and  $WTP$  for buying a lottery via the BDM mechanism. Hence, it could be the case that embedding our experiment into the legal context increases the gap in comparison to earlier abstract trades of lotteries. Yet, this increase could be also attributed to the fact that we ask for subjects'  $WTA$  for accepting the ownership of lottery and their  $WTP$  for buying into a safe payoff, while previous studies analyzed the contrary case.

With regard to the hypotheses in section 3.3, the non-parametric tests affirm Hypothesis 1. There is no indication for Hypothesis 2 that the  $WTA/WTP$  disparity decreases over the course of rounds. This leaves room for the interpretation that the gap is not driven by misconceptions about the elicitation device: subjects do not learn over time.<sup>27</sup> However, in addition

---

<sup>25</sup>Table 3 also indicates that there are fewer observations for  $WTA/WTP$  than for  $WTA$  and  $WTP$ . Through the standardization, all observations with  $WTP = 0$  are lost. However, the power of the hypothesis test is not impaired, since the Wilcoxon test compares the two matched  $WTA$  and  $WTP$  samples and does not rely on the measured gap. The gap serves as additional information that can be compared with other experiments.

<sup>26</sup>In their article, Horowitz and McConnell (2002, p. 433) report mean ratios for different types of good. We calculated the median ratios on the basis of the individual median gaps listed in the appendix using the same method as Horowitz and McConnell (2002, p. 433). The mean ratio for lotteries is 2.10 and for ordinary private goods is 2.92.

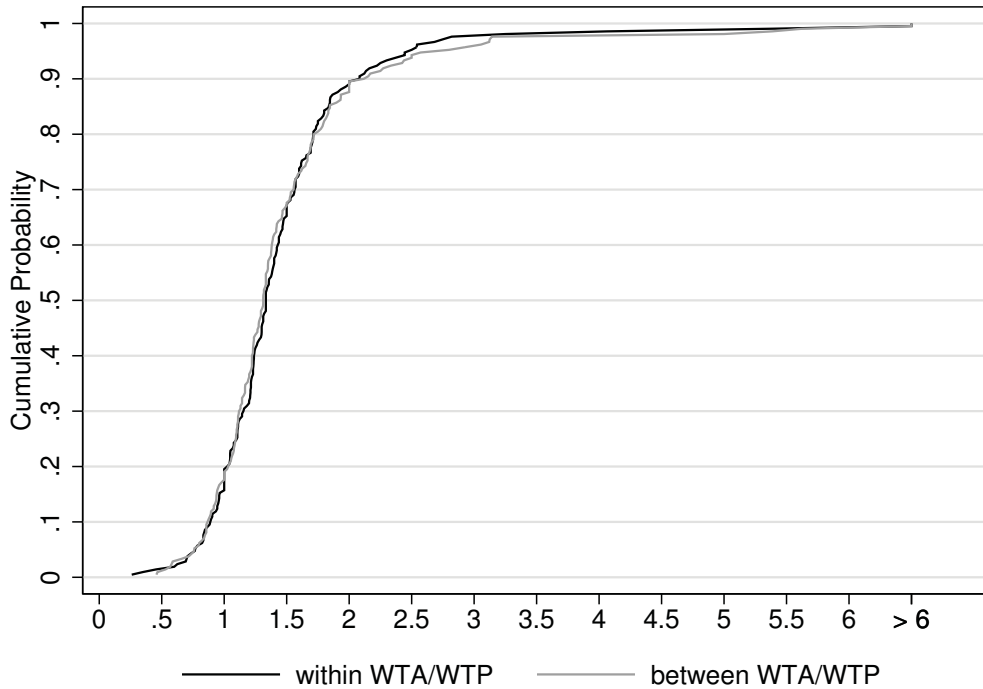
<sup>27</sup>This result stands at odds with other literature, such as Coursey et al. (1987) and Plott and Zeiler (2005), who emphasize the ability to learn.

we vary low-stake and high-stake contract pairs over rounds, so that we will come back to this point in the regression analysis, which allows us a more precise examination of Hypothesis 2.

## 4.2 Potential Caveats

This section will control for potential caveats. First, it will be explored if it matters whether the  $WTA/WTP$  disparity is constructed as a within gap or a between gap. Second, it will be investigated if the observed  $WTA/WTP$  disparity is caused by irrational valuations, which might weaken the power of the previously found results.

Figure 2: Within and between  $WTA/WTP$



*Baseline* allows us to observe between-subjects'  $WTA/WTP$  ratios along within-subjects'  $WTA/WTP$  ratios. Again, also in a between-subjects comparison for almost all contract pairs,  $WTA$  significantly exceeds  $WTP$  at the 1 or 5 percent level (except rounds 6 and 7).<sup>28</sup> Figure 2 compares

<sup>28</sup>The table 6 presenting the summary statistics of the between-subjects  $WTA/WTP$

the within-subjects' and between-subjects'  $WTA/WTP$  and shows their cumulative probability distribution.<sup>29</sup> The distributions seem very similar and also a Wilcoxon signed rank test cannot reject the hypothesis that the within-subjects' gap is equal to the between-subjects' gap of the corresponding contract pair. Thus, we cannot confirm the results of the meta-study by Sayman and Öncüler (2005) that show a significant difference between within-subjects' and between-subjects'  $WTA/WTP$  gaps. They argue that, in contrast to between designs, in within designs individuals might realize that  $WTP$  and  $WTP$  tasks technically pose the same question and for the sake of consistency offer the same responses. Our data is not indicative of this problem.

Plott and Zeiler (2011, p. 1018-1020) find a substantial number of irrational valuations in their lottery data. These irrational valuations lie outside the bounds of the lottery value, which means that irrational valuations will create losses for the subject if a contract exchange takes place. Irrational valuations are problematic because they might indicate that subjects do not understand the task, the concepts of randomization or probabilities. Therefore, our data will also control for irrational valuations.

While Plott and Zeiler define irrational valuations as those that fall outside the support of a lottery value's probability distribution (Plott and Zeiler, 2011, p. 1020),<sup>30</sup> we define irrational valuations according to values above  $Max$ ,<sup>31</sup> since our subjects need to compare two payoffs when making their valuation. By design, such irrational valuations could only be observed for the contract pairs 3, 6, and 7. For all other contract pairs, the maximal valuation is within the rational range as all possible valuations are limited by the endowment given at the beginning of each round.

Notice  $Max$  is only an imperfect criterion for irrationality of  $WTA$  in *baseline*. That is, valuations higher than  $Max$  in this treatment could also

---

ratio, called in short *between WTA/WTP gap*, can be found in the Appendix A. Since between gaps are based on  $WTA$  and  $WTP$  valuations elicited in the same period, these gaps are not averaged over the whole matching group, but only over each subgroup (a subgroup contains three player 1 or three player 2 subjects per group). Thus, for each contract pair, a between gap equals the mean of three  $WTA$  valuations (belonging to one subgroup) divided by the mean of three  $WTP$  valuations (from the other subgroup) elicited in the same period. Therefore, in each group the between-subjects  $WTA/WTP$  ratio is elicited twice.

<sup>29</sup>Gaps from all subjects over all rounds in *baseline* are included in the figure. To have a meaningful comparison with the between gaps, the within gaps are also constructed on the basis of the subgroup.

<sup>30</sup>For instance, if a lottery has a 50 percent probability of paying  $x$  and a 50 percent probability  $y$  with  $y < x$ , then valuations above  $x$  or below  $y$  violate rationality postulates.

<sup>31</sup>We explained the rationale behind  $Max$  in the footnote 20 of section 3.1.

result from strategic incentives of the elicitation mechanism paired with extreme risk aversion (i.e., if subjects are very risk-averse, so that their individual  $WTA$  almost equals  $X - y$ ,  $WTA$  could exceed  $Max$ ). Therefore, we potentially overestimate results on irrationality based on  $WTA$  in *baseline*. However, since there is little difference between the percentages of irrationality between *baseline* and *variation*, particularly for the numbers based on  $WTA$ , we conclude that this effect is rather minor.

Figure 3: Percentage of irrational valuations

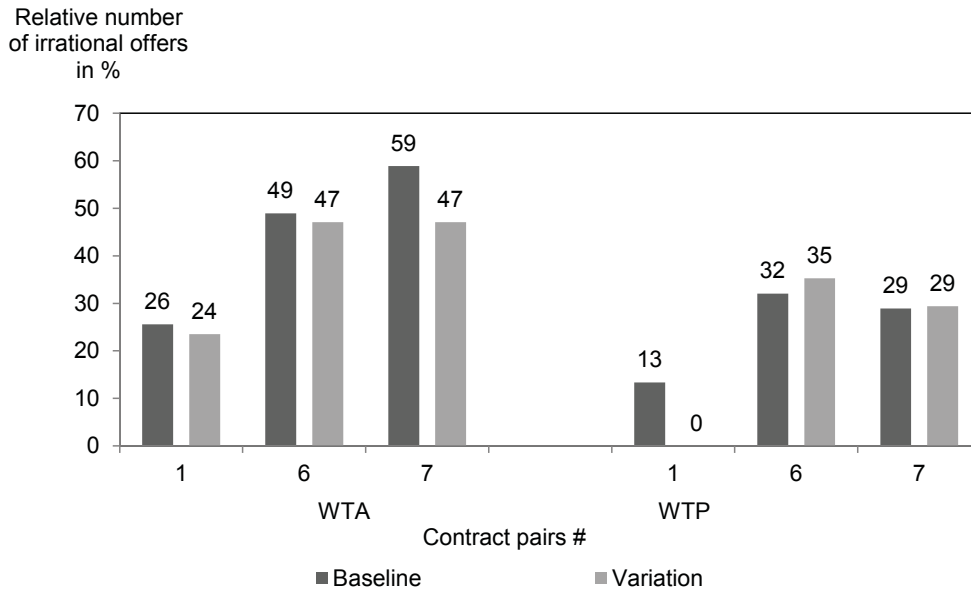


Figure 3 shows that in some lotteries large fractions of valuations are irrational.<sup>32</sup> For  $WTA$ , the percentages of irrational offers lie between 24% and 59% for the contract pairs in question. For  $WTP$ , the numbers are substantially lower and range between 0 and 35%. While irrational valuations are also found in Plott and Zeiler (2011) and Isoni et al. (2011), their levels are lower than ours. This may be related to the fact that bargaining experiments are rarely run at the laboratory for economics and social science at the University of Hamburg, and that this type of experiment is completely new to all participants. The disparity of irrational offers between the low-stake contract pair 3 and the high-stake contract pairs 6 and 7 may be explained by the different opportunities of making an irrational offer. In the high-stake

<sup>32</sup>Figure 3 shows irrational valuations on an individual basis.

tasks, subjects dispose of a higher endowment, which increases the opportunity to make an irrational offer by three times compared to the low-stake tasks.<sup>33</sup> The crucial question is whether the large degree of irrationality simply causes the *WTA/WTP* gap (particularly since irrationality affects *WTA* more likely than *WTP*).

Table 4 shows *WTA/WTP* gaps for *baseline* as well as *variation* excluding irrational valuations.<sup>34</sup> For *baseline* the significance of the results is not changed, but for *variation* it is substantially reduced (yet only for one contract pair does the gap turn out to be insignificant). The results show that the endowment effect is robust for *baseline*, but partly disappears for *variation* once we eliminate irrational offers. This result supports the claim that the gap is partly driven by irrational valuations. However, even if misconceptions are the driving force of irrational valuations, they do not account fully for the observed gap.

To explore whether the different elicitation methods of *baseline* and *variation* result in different valuations, a two-tailed Mann-Whitney ranksum test examines the hypothesis that the valuations under *baseline* are equal to the valuations under *variation*.<sup>35</sup> Surprisingly, there are only unsystematic differences between the treatment conditions. We find a higher average WTP in round eight of *baseline*, compared with *variation*, at a significance level of one percent. All other differences – particularly the ones for WTA – are not significantly different at a ten-percent level. Therefore, it seems that the constant change between the incentive-incompatible WTA and incentive-compatible WTP elicitation in *baseline* eliminates systematic treatment differences. Rather, as we will show in the subsequent section, another factor – the differences in expected values – crucially influences differences between WTA and WTP under both treatment conditions.

The results of the descriptive statistics and non-parametric tests highlight that over all rounds and treatments significant *WTA/WTP* gaps are found. Thus, it can be concluded that Hypothesis 1 is affirmed, while we cannot support Hypothesis 2, although it seems that misconceptions amplify the endowment effect. Finally, Hypothesis 3 is tested in the subsequent section by means of a regression analysis.

---

<sup>33</sup>For the low-stake contract pairs, the endowment amounts to 0.5 ECU and an offer is defined as irrational if it is greater than 0.4 ECU. For the large stake contract pairs, where the endowment amounts to 5 ECU, an offer is defined as irrational if the offer is greater than 2 ECU.

<sup>34</sup>The *WTA/WTA* gaps are based on matching group mean valuations for *baseline*. When constructing the mean, irrational individual valuations were excluded.

<sup>35</sup>Again, we use matching group means for *baseline* and individual valuations for *variation*.

Table 4: Summary of  $WTA/WTP$  without irrational valuations

Contract pairs	3	6	7
<i>baseline:</i>			
	<i>WTA</i>		
n	15	14	13
Mean	0.29	1.39	1.50
Median	0.30	1.45	1.43
SD	0.05	0.40	0.38
	<i>WTP</i>		
n	15	15	15
Mean	0.25	1.04	1.14
Median	0.26	1.00	1.01
SD	0.05	0.39	0.31
Significance: $WTA=WTP^1$	***	***	***
	<i>WTA/WTP</i>		
n	15	14	13
Mean	1.17	1.46	1.44
Median	1.18	1.30	1.40
<i>variation:</i>			
	<i>WTA</i>		
n	13	9	9
Mean	0.28	0.77	1.06
Median	0.25	0.8	1
SD	0.08	0.81	0.94
	<i>WTP</i>		
n	17	11	12
Mean	0.24	0.62	0.7
Median	0.25	0.5	0.5
SD	0.11	0.73	0.72
Significance: $WTA=WTP^1$		*	**
	<i>WTA/WTP</i>		
n	13	4	5
Mean	1.23	1.30	4.40
Median	1.2	1.5	4

<sup>1</sup> Test based on signed ranks (two-tailed Wilcoxon signed rank test).

Note: Significance levels are indicated as \*\*\* (1%), \*\* (5%), and \* (10%).



### 4.3 Regression Analysis

The regression analysis aims to identify the influence of explanatory variables on the dependent variable, the *WTA/WTP* gap. The experimental data consist of data points for 107 subjects over 14 periods that can be classified as a balanced panel. In *baseline*, subjects are matched with other subjects and remain in the same group for the entire duration of the experiment, so that the regression error terms might be correlated within matching groups. A mixed effects model that allows modeling random effects on the group and on the individual level will therefore be applied. In total, there are 32 groups composed of 15 matching groups from *baseline*, and 17 groups from *variation* where each group contains one subject.

In the mixed effects model, the error term  $v_{it}$  consists of three components:  $c_k$  is the random intercept for group  $k$ ,  $\epsilon_{ik}$  is the random intercept for group  $k$  and subject  $i$ , and  $u_{ikt}$  is the random effect for group  $k$ , subject  $i$  at time period  $t$ . The random effect for the group is nested within subjects, which means that it takes on a different value for each combination of group and subject (Rabe-Hesketh and Skrondal, 2005, pp. 436-437).

Thus, the mixed effects model can be written as

$$WTA/WTP_{ikt} = X_{it}\beta + c_k + \epsilon_{ik} + u_{ikt} \quad (9)$$

where  $X_{it}$  denotes the matrix of regressors. In our regressions, we face the following problem: WTA and WTP values for each lottery pair are elicited in different periods, which prohibits a direct assignment of the gap to one period. Therefore, the gap will be assigned to the period where the first value of the gap was elicited (variable *period*). Additionally, the absolute difference of periods between the WTA and WTP measurement will be captured in the variable *distance*, defined as  $distance = |period_{WTA} - period_{WTP}|$ .<sup>36</sup> In addition, we include the dummy variable *treatment* (which equals 0 for observations from *baseline* and 1 from *variation*),  $\Delta EV$  (the difference between the expected values of the lotteries included in contract A and B), and a dummy variable, *highstake* (which equals 0 for all low-stake rounds and 1 for all high-stake lotteries). Furthermore, interaction terms between *period* and *treatment* (*period\_treat*), *distance* and *treatment* (*distance\_treat*), *highstake* and *treatment* (*highstake\_treat*),  $\Delta EV$  and *treatment* ( $\Delta EV\_treat$ ) are added to the specification. Table 5 reports the results for regressions (1)-

---

<sup>36</sup>Thus, *period* runs from 1 to 13, while it never takes the value 14, as the last period always elicits the second value of the gap. *Distance* runs from 1 to 7, as the maximal distance between WTA and WTP periods within a block (lowstake/highstake) is seven periods.

(5) and a joint test for the goodness of fit for the complete model (Wald test).

Table 5: Mixed effects regression

	(1)	(2)	(3)	(4)	(5)
<i>treatment</i>	0.294 (0.562)	0.300 (0.561)	0.304 (0.561)	0.385 (0.561)	1.121 (1.738)
<i>highstake</i>		0.908** (0.368)	1.418*** (0.430)	-0.0564 (0.977)	-0.219 (1.073)
$\Delta EV$			-0.437** (0.193)	-0.512** (0.203)	-0.445** (0.218)
<i>period</i>				0.223* (0.127)	0.229 (0.140)
<i>distance</i>				0.253* (0.153)	0.289* (0.164)
<i>highstake_treat</i>					1.302 (2.754)
$\Delta EV\_treat$					-0.454 (0.631)
<i>period_treat</i>					-0.0804 (0.340)
<i>distance_treat</i>					-0.253 (0.472)
<i>constant</i>	1.950*** (0.218)	1.433*** (0.302)	1.502*** (0.303)	0.348 (0.652)	0.255 (0.702)
Observations	688	688	688	688	688
Number of groups	32	32	32	32	32
Wald Test	0.27	6.36**	11.57***	15.72***	17.21**

Notes: Standard errors in parentheses; significance at the 10, 5 and 1 percent level is denoted by \*, \*\* and \*\*\*, respectively.

The coefficients of the variable *treatment* as well as the ones of the interaction terms turn out to be insignificant in all regressions, which indicate that the different value elicitation methods applied in *baseline* and *variation* do not significantly alter the gap. This result may be surprising given the incentives for the truthful revelation of valuations in *baseline* and *variation*, but adds to the observations of other authors indicating that the BDM mech-

anism leads to behavior which does not necessarily correspond to theoretical predictions (Noussair et al., 2004).

The coefficients of variable *highstake* in (2) and (3) suggest a positive influence of high-stake lotteries on the size of the gap. However, this effect disappears once we control for timing of the elicitation. The coefficient for the variable *period* is positive and significant in model (4) at the ten-percent level, which (weakly) signifies that the later the valuations were elicited, the higher the gap is. However, the variable turns insignificant once the interaction terms are included. More importantly, the coefficient of the variable *distance* is significant in regressions (4) and (5) at a ten-percent level. This indicates that the more periods have passed between both WTA and WTP measurements, the higher the gap is. This result affirms the finding by Sayman and Öncüler (2005) that subjects try to give consistent valuations. The more distant the valuations are in time, the more the subjects may forget about their previous valuation. Hence, we cannot confirm Hypothesis 2 as the gap does not decrease over time. This result stands in contrast to the positive effect of experience found by Plott and Zeiler (2005). However, it is in line with results of a meta-study by (Horowitz and McConnell, 2002, p. 427), who find mixed effects of repetition on the WTA-WTP disparity.

Most importantly, the variable  $\Delta EV$  exerts the strongest impact on the *WTA/WTP* gap. Its coefficient is negative and significant at a five-percent level and affects consistently the gap in models (3)-(5) by roughly  $-0.5$ . The negative sign demonstrates that a high difference in expected values between the lotteries reduces the size of the gap. This result affirms Hypothesis 3: the more the alternative rule differs from the default rule, the less subjects' valuations diverge. This result points out that subjects may act more rationally for higher differences in expected value.

Overall, the non-parametric tests already confirmed Hypothesis 1 by detecting a statistical significant difference between WTA and WTP valuations. The regression analysis suggests the different elicitation mechanisms do not produce significantly different *WTA/WTP* gaps. Hypothesis 2 assumes that the gap decreases over time. We found no indication for this hypothesis. On the other hand, we do not have consistent evidence for an increase of the gap over periods. Thus, neither we can support the argumentation that the gap is driven by misconceptions, nor do the results suggest that the gap disappears due to subject's learning. Hypothesis 3, however, is affirmed by the regression analysis. The differences in expected value turn out to be significant in the mixed effects model. It appears that the more both contracts differ in terms of expected payoff, the less subjects are affected by the endowment effect.

## 5 Conclusion

This paper has shown that default rules are more powerful than their theoretical justification suggests at first sight. Created for contract law simply to address gaps in incomplete contracts, default rules do create distortions at closer inspection. This paper has used an experiment to show that the default rule creates an endowment effect: the minimal WTA for a loss of the default rule was found to be greater than the maximal WTP to obtain the same rule when it is not set as default.

Our experimental design takes into account previous methodological concerns (e.g., practice rounds) raised by economic scholars (Plott and Zeiler, 2005, 2007). In contrast to previous experiments in law (Korobkin, 1998b; Sunstein, 2002), therefore, we can show that the gap is not driven by a lack of experimental controls. On the other hand, we test the endowment effect for default rules in a legal setting, thereby also addressing the legal critique (Tor, 2007) that stresses the importance of introducing a legal context. Particularly, modeling the liability rule under scrutiny as a lottery provides two crucial advantages: first, the results of our study can be compared to WTA-WTP studies on lotteries, so that they add to the general literature on the endowment effect for goods of uncertain value. Second, the legal setting was still sufficiently abstract in order not to trigger any reference to a particular legal system, which could have evoked other psychological biases such as the status quo bias.

The gap between WTA and WTP was statistically significant and persistent over time. The median gap amounts to 1.31 in *baseline* and 1.45 in *variation*. Thus, the more realistic baseline treatment – albeit not incentive-compatible for WTA valuations – induces a lower gap. In comparison to other studies that test the endowment effect for lotteries, we find gaps of similar, yet larger size (e.g., Plott and Zeiler, 2005, report medians of 1.15-1.16). Partly, this could result from the legal context of our experimental setting. Moreover, additional robustness checks are conducted. First, the valuations are controlled for irrationality. Although a substantial percentage of irrational gaps is found, the results show that irrationality is not the driving force of the gap: even in the sub-sample of only rational valuations, there is a significant *WTA/WTP* gap. Second, gaps constructed as between or within gaps do not significantly differ. This result emphasizes that an endowment effect cannot only be observed within individuals, but also seems to persist over a population average. Third, a high difference in expected values between the lotteries persistently reduces the size of the gap.

What are legal consequences of our results? Default rules are not innocent. The experimental results emphasize the need for a theory of an

optimal design of default rules.<sup>37</sup> Bechtold (2010) claims that the majority of experimental studies justify the construction of default rules according to the majoritarian opinion because subjects opt out of default rules less often than predicted. However, our experiment also points out that default rules affect stated valuations, which leads to the problem that, under different default rules, agents agree on different contract terms. Therefore, it might be problematic to base the optimal design of default rules only on the revealed interests of agents; a normative justification of a default rule may also be necessary.

Independently from the issue of optimal default rules, our results suggest a potential way to reduce the biased valuation of default rights due to endowment effects: our results show a consistent decrease of the *WTA/WTP* gaps for differences of expected values in the underlying contracts. As mentioned earlier, we interpret this finding in a way that subjects are more likely to make a high effort when elaborating their valuations for higher salience. In other words, the gap decreases if subjects work hard on their *WTA* or *WTP* statement. This may serve as an opportunity for lawyers and policy-makers. If one urges contracting parties to agree or to justify default clauses explicitly, this may also increase the effort parties put into the contract thereby hopefully reducing the influence of psychological biases such as the endowment effect. The proof of this claim, however, lies beyond the scope of this paper and should be a promising field of further research.

## References

- Adler, B. E. (1999). The questionable ascent of *Hadley v. Baxendale*. Stanford Law Review 51(6), pp. 1547–1589.
- Arlen, J., M. Spitzer, and E. Talley (2002). Endowment effects within corporate agency relationships. The Journal of Legal Studies 31(1), 1–37.
- Arlen, J. and S. Tontrup (2013). A process account of the endowment effect: Voluntary debiasing through agents and markets. New York University Public Law and Legal Theory Working Papers 399.
- Ayres, I. and R. Gertner (1989). Filling gaps in incomplete contracts: An economic theory of default rules. The Yale Law Journal 99(1), 87–130.

---

<sup>37</sup>We follow the same argumentation as Camerer and Talley (2007, p. 1633). They argue that more research has to deal with the question of how to set default rules in contracting if an endowment effect exists in legal frameworks.

- Bardsley, N., R. Cubitt, G. Loomes, P. S. C. Moffatt, and R. Sugden (2009). Experimental Economics: Rethinking the Rules. Princeton University Press.
- Bebchuk, L. and S. Shavell (1991). Information and the scope of liability for breach of contract: The rule of Hadley v. Baxendale. Journal of Law, Economics, & Organization 7(2), 284–312.
- Bechtold, S. (2010). Die Grenzen zwingenden Vertragsrechts: Ein rechtsökonomischer Beitrag zu einer Rechtsetzungslehre des Privatrechts. Mohr Siebeck GmbH KG.
- Becker, G. D. M. M. J. (1964). Measuring utility by a single-response sequential method. Behavioral Science 9(3), 226–236.
- Bernstein, L. (1993). Social norms and default rules analysis. Southern California Interdisciplinary Law Journal 3, 59–91.
- Bock, O., A. Nicklisch, and I. Baetge (2012). hroot: Hamburg registration and organization online tool. H-Lab Working Paper (1).
- Buccafusco, C. and C. Sprigman (2011). The creativity effect. The University of Chicago Law Review 78(1), 31–52.
- Camerer, C., G. Loewenstein, and M. Rabin (2003). Advances in behavioral economics. Princeton University Press.
- Camerer, C. and E. Talley (2007). Chapter 21 experimental study of law. In A. Polinsky and S. Shavell (Eds.), Handbook of Law and Economics, Volume 2, pp. 1619–1650. Elsevier.
- Carroll, G. D., J. J. Choi, D. Laibson, B. Madrian, and A. Metrick (2009). Optimal defaults and active decisions. The Quarterly Journal of Economics 124(4), 1639–1674.
- Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2004). For better or for worse: Default effects and 401 (k) savings behavior. In D. A. Wise (Ed.), Perspectives on the Economics of Aging, pp. 81–126. University of Chicago Press.
- Coursey, D. L., J. L. Hovis, and W. D. Schulze (1987). The disparity between willingness to accept and willingness to pay measures of value. The Quarterly Journal of Economics 102(3), 679–690.

- Eisenberger, R. and M. Weber (1995). Willingness-to-pay and willingness-to-accept for risky and ambiguous lotteries. Journal of Risk and Uncertainty 10(3), 223–233.
- Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10(2), 171–178.
- Geis, G. (2005). Empirically assessing Hadley v. Baxendale. American Law & Economics Association Annual Meetings, 897–956.
- Harless, D. W. (1989). More laboratory evidence on the disparity between willingness to pay and compensation demanded. Journal of Economic Behavior & Organization 11(3), 359–379.
- Horowitz, J. K. and K. E. McConnell (2002). A review of WTA/WTP studies. Journal of Environmental Economics and Management 44(3), 426–447.
- Isoni, A., G. Loomes, and R. Sugden (2011). The willingness to pay–willingness to accept gap, the endowment effect, subject misconceptions, and experimental procedures for eliciting valuations: Comment. The American Economic Review 101(2), 991–1011.
- Johnson, E. and D. Goldstein (2003). Do defaults save lives? Science 302(5649), 1338–1339.
- Johnson, E., J. Hershey, J. Meszaros, and H. Kunreuther (1993). Framing, probability distortions, and insurance decisions. Journal of Risk and Uncertainty 7(1), 35–51.
- Kahan, M. and M. Klausner (1997). Standardization and innovation in corporate contracting (or the economics of boilerplate). Virginia Law Review 83(4), 713–770.
- Kahneman, D., J. Knetsch, and R. Thaler (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. The Journal of Economic Perspectives 5(1), 193–206.
- Kahneman, D., J. L. Knetsch, and R. Thaler (1986). Fairness as a constraint on profit seeking: Entitlements in the market. American Economic Review 76(4), 728–741.
- Karni, E. and Z. Safra (1987). Preference reversal and the observability of preferences by experimental methods. Econometrica 55(3), 675–685.



- Knetsch, J. (1989). The endowment effect and evidence of nonreversible indifference curves. The American Economic Review 79(5), 1277–1284.
- Knetsch, J., F.-F. Tang, and R. Thaler (2001). The endowment effect and repeated market trials: Is the vickrey auction demand revealing? Experimental Economics 4(3), 257–269.
- Knetsch, J. and W. Wong (2009). The endowment effect and the reference state: Evidence and manipulations. Journal of Economic Behavior & Organization 71(2), 407–413.
- Korobkin, R. (1998a). Inertia and preference in contract negotiation: The psychological power of default rules and form terms. Vanderbilt Law Review 51, 1583–1651.
- Korobkin, R. (1998b). Status quo bias and contract default rules. Cornell Law Review 83, 608–687.
- Kovalchik, S., C. Camerer, D. Grether, C. Plott, and J. Allman (2005). Aging and decision making: A comparison between neurologically healthy elderly and young individuals. Journal of Economic Behavior & Organization 58(1), 79–94.
- List, J. (2003). Does market experience eliminate market anomalies? The Quarterly Journal of Economics 118(1), 41–71.
- List, J. A. and C. A. Gallet (2001). What experimental protocol influence disparities between actual and hypothetical stated values? Environmental and Resource Economics 20(3), 241–254.
- Madrian, B. and D. Shea (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. The Quarterly Journal of Economics 116(4), 1149–1187.
- Milgrom, P. (2004). Putting auction theory to work. Cambridge University Press.
- Noussair, C., S. Robin, and B. Ruffieux (2004). Revealing consumers’ willingness-to-pay: A comparison of the BDM mechanism and the Vickrey auction. Journal of Economic Psychology 25(6), 725–741.
- Peters, E., P. Slovic, and R. Gregory (2003). The role of affect in the WTA/WTP disparity. Journal of Behavioral Decision Making 16(4), 309–330.

- Plott, C. and K. Zeiler (2005). The willingness to pay–willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations. The American Economic Review 95(3), 530–545.
- Plott, C. and K. Zeiler (2011). The willingness to pay–willingness to accept gap, the "endowment effect," subject misconceptions, and experimental procedures for eliciting valuations: Reply. The American Economic Review 101(2), 1012–1028.
- Plott, C. R. and K. Zeiler (2007). Exchange asymmetries incorrectly interpreted as evidence of endowment effect theory and prospect theory? American Economic Review 97(4), 1449–1466.
- Posavac, S. (1998). Strategic overbidding in contingent valuation: Stated economic value of public goods varies according to consumers expectations of funding source. Journal of Economic Psychology 19(2), 205–214.
- Rabe-Hesketh, S. and A. Skrondal (2005). Multilevel and longitudinal modeling using Stata. College Station and Tex: Stata Press.
- Sayman, S. and A. Öncüler (2005). Effects of study design characteristics on the WTA–WTP disparity: A meta analytical framework. Journal of Economic Psychology 26(2), 289–312.
- Schmidt, U. and S. Traub (2009). An experimental investigation of the disparity between WTA and WTP for lotteries. Theory and Decision 66(3), 229–262.
- Schwab, S. (1988). A Coasean experiment on contract presumptions. The Journal of Legal Studies 17(2), 237–268.
- Shogren, J., S. Shin, D. Hayes, and J. Kliebenstein (1994). Resolving differences in willingness to pay and willingness to accept. The American Economic Review 84(1), 255–270.
- Sloof, R., H. Oosterbeek, and J. Sonnemans (2007). On the importance of default breach remedies. Journal of Institutional and Theoretical Economics 163(1), 5–22.
- Spier, K. (1992). Incomplete contracts and signalling. The RAND Journal of Economics 23(3), 432–443.
- Sugden, R., J. Zheng, and D. J. Zizzo (2013). Not all anchors are created equal. mimeo.

- Sunstein, C. (2002). Switching the default rule. New York University Law Review 77, 106–134.
- Thaler, R. (1980). Toward a positive theory of consumer choice. Journal of Economic Behavior & Organization 1(1), 39–60.
- Tor, A. (2007). On contractual defaults and experimental law and economics. Journal of Institutional and Theoretical Economics 163(1), 26–29.
- Trautmann, S. T. and U. Schmidt (2012). Pricing risk and ambiguity: the effect of perspective taking. Quarterly Journal of Experimental Psychology 65(1), 195–205.

## A Additional Tables

Table 6: Summary of between  $WTA/WTP$  ratios for *baseline*

Contract pairs	1		2		3		4		5		6		7	
	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
n	15	15	14	15	15	15	15	15	15	15	15	15	15	15
Mean	1.39	1.08	1.43	1.60	1.18	1.38	1.38	1.58	1.41	1.31	1.48	2.95	1.56	1.74
Median	1.22	1.13	1.37	1.38	1.09	1.33	1.33	1.29	1.47	1.28	1.42	1.38	1.42	1.75
Significance: $WTA=WTP^1$	***		**	***		***	***	***	***	***	**	**	***	***

<sup>1</sup> Test based on two-tailed Wilcoxon rank-sum test.

Note: Significance levels are indicated as \*\*\* (1%), \*\* (5%), and \* (10%).

1st stands for the first time the between  $WTA/WTP$  ratio was elicited in each group and 2nd for the second time, respectively.

## **B English Translation of the Instructions for the Baseline Treatment**

On the following pages, the English translation of the instructions for the baseline treatment is reproduced.

## I. General Instructions for Participants

Welcome to the experiment!

You are now taking part in an economic experiment. You will be able to earn a considerable amount of money, depending on your decisions and the decisions of others. It is therefore important that you read these instructions carefully.

The instructions which we have distributed to you are solely for your private information. It is prohibited to communicate with other participants during the experiment. Please do not respond verbally to the events that occur during the experiment. Should you have any questions, please raise your hand and an experimenter will come to answer them. If you violate this rule, we will have to exclude you from the experiment and from all payments.

During the experiment you will make decisions anonymously. Only the experimenter knows your identity while your personal information is confidential and your decisions will not be traceable to your identity.

For your timely arrival and participation in this experiment you will earn 5 Euros. The additional earnings depend on your decisions. During the experiment your earnings will be calculated in tokens. At the end of the experiment your earned tokens will be converted into Euro at the following exchange rate:

**4 Tokens = 1 EURO,**

and will be paid to you in cash. Please wait in your cabin until you are asked to collect your payment. Please bring all documents you have received from us.

This experiment examines factors of buying and selling decisions. The experiment consists of 17 periods. The exact course of a round will be explained in the following.

## II. Background information

In this experiment, each participant represents a company or a customer. The computer will show you whether you are a company or a customer. This assignment is valid for the entire experiment.

The company specializes in the supply of goods and is located in the state A. The customer is located in country B. The company and the customer have agreed on general terms on the delivery of goods. However, for each contract it must be specified how the liability for consequential costs in case of late deliveries will be determined, since the countries A and B have different rules governing the liability for late delivery. Now, the company and the buyer agree under which of the two liability rules the contract comes into force.

### **Rule A from country A:**

*Suppliers who deliver later than agreed **are not held liable** for the resulting costs.*

The **company** is therefore not held liable and receives the **same** payment in both events, on-time delivery or late delivery. The **customer**, in contrast, receives payments of **different** sizes, depending on the time of delivery.

#### Rule B from country B:

*Suppliers who deliver later than agreed **are held liable** for the resulting costs.*

The **company** is therefore held liable for late delivery and, depending on the time of delivery, receives payments of **different** sizes. **The customer**, in contrast, receives the **same** payment in both events, on-time delivery or late delivery.

The probability of an on-time delivery as well as the probability of a late delivery is 50%. On average, **50 out of 100** deliveries are **on time**, while 50 out of 100 deliveries arrive **late**.

### III. Information on the course of a round

You will receive an endowment in each round and you face a contract that follows either rule A or B. You can spend part of your tokens for switching to the other rule or you can receive tokens for accepting the other rule. For this purpose, the company or the customer states the price that they are willing to spend for a rule change. At the same time, the company or the customer indicates at what price they are willing to accept such a change.

#### Example 1: The contract follows rule A

You get told how much you can earn under the contract. Four different cases with different contract payments can occur, which differ depending on the applied rule and the time of delivery. The first number identifies the payment for the company; the second number is the payment for the customer:

	On-time delivery - 50 out of 100 cases -	Late delivery - 50 out of 100 cases -
Rule A	5 / 5 Tokens	5 / 4 Tokens
Rule B	5 / 5 Tokens	4 / 5 Tokens

The customer specifies the **highest amount – the maximum purchase price** – which it is willing to pay for a change to rule B (renegotiations are not possible). Simultaneously, the company states at what **minimum amount – the minimum selling price** – he would be willing to accept a change to rule B (renegotiations are not possible).

The prices indicated by the company and the customer will be compared. If **the minimum selling price** is less than or equal to the **maximum purchase price**, both switch to Rule B. The company receives the minimum selling price, and the customer pays the minimum selling price. The contract payments now follow rule B. The company receives the contract payment plus the endowment plus the minimum selling price. The customer receives the contract payment plus the remaining endowment minus the minimum selling price.



If the **minimum selling price** is greater than the **maximum purchase price**, there will be no contract change and the contract payment will continue to follow rule A. The company receives the contract payment plus the initial endowment. The customer receives the contract payment plus the initial endowment.

Numerical example:

*If the company indicates 0.70 tokens as its minimum selling price and the customer specifies 0.80 tokens as his maximum purchase price, the rule will be exchanged. The company receives 0.70 tokens and the customer pays 0.70 tokens. Both receive the contract payment following rule B.*

*If the company indicates 1.00 tokens as its minimum selling price and the customer specifies 0.80 tokens as his maximum purchase price, there is no change since the minimum selling price is higher than the maximum purchase price. Rule A will still be applied and both receive the contract payment following rule A.*

## **Example 2: The contract follows rule A**

The company now specifies the **highest amount – the maximum purchase price** – which it is willing to pay for a change to rule A (renegotiations are not possible). Simultaneously, the customer states at what **minimum amount – the minimum selling price** – he would be willing to accept a change to rule A (renegotiations are not possible).

The prices indicated by the company and the customer will be compared. If the **minimum selling price** is less than or equal to the **maximum purchase price**, both switch to Rule A. The customer receives the minimum selling price, and the company pays the minimum selling price. The contract payments now follow rule A. The customer receives the contract payment plus the endowment plus the minimum selling price. The company receives the contract payment plus the remaining endowment minus the minimum selling price.

If the **minimum selling price** is greater than the **maximum purchase price**, there will be no contract change and the contract payment will continue to follow rule B. The company receives the contract payment plus the initial endowment. The customer receives the contract payment plus the initial endowment.

Numerical example:

*If the customer indicates 0.70 tokens as his minimum selling price and the company specifies 0.80 tokens as its maximum purchase price, the rule will be exchanged. The customer receives 0.70 tokens and the company pays 0.70 tokens. Both receive the contract payment following rule A.*

*If the customer indicates 0.70 tokens as his minimum selling price and the company specifies 0.50 tokens as its maximum purchase price, there is no change since the minimum selling price is higher than the maximum purchase price. Rule B will still be applied and both receive the contract payment following rule B.*

#### IV. Information on the course of the experiment

The whole experiment consists of the following five parts. The possible income from previous rounds cannot be carried over to the subsequent rounds.

- Part 1) Three practice rounds: In every practice round, you receive a contract for a delivery of low value. For each contract you will receive an endowment of 0.50 tokens. The practice rounds aim to give you an understanding of the rules. They do not affect your final payment.
- Part 2) Six rounds of low value deliveries: In every practice round, you receive a contract for a delivery of low value. For each contract you will receive an endowment of 0.50 tokens.
- Part 3) Eight rounds of high value deliveries: In every practice round, you receive a contract for a delivery of high value. For each contract you will receive an endowment of 5 tokens.
- Part 4) Closure: After the experiment, you will be asked to fill out a questionnaire on general issues.
- Part 5) Payout: You will receive your earnings (contract payment (+/- transaction price in case of contract change) + remaining endowment) of a randomly selected round of low delivery value and of a randomly selected round of high delivery value (converted to Euro), but not the earnings from the practice rounds, and 5 Euro paid for your timely appearance in cash.

In each round, you are randomly paired with a customer or with a company. The contracts in each round are independent from each other. The contracts within the parts 1, 2 and 3 are assigned at random. The structure of the contracts is identical except for the different levels of payments. In each round, you get a contract that either follows rule A or B.

If you have questions about the experiment, please raise your hand now.