Does the Fundamental Transformation Deter Trade? An Experiment

Christoph Engel, *Max Planck Institute for Research on Collective Goods,* and Eric Helland, *Claremont McKenna University*

Send correspondence to: Christoph Engel, Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Straße 10, 53113 Bonn, Germany; E-mail: engel@coll.mpg.de

Oliver Williamson has coined the term "fundamental transformation." It captures the following situation: before they strike a deal, buyers and sellers are protected by competition. Yet, thereafter, they find themselves in a bilateral monopoly. With common knowledge of standard preferences, both sides would conclude the contract regardless if its expected value exceeds their outside options. We run an experiment to test whether additional behavioral concerns deter mutually beneficial trade. We test four concerns: If the risk materializes, another individual makes a windfall profit; she does so by intentionally exploiting another individual; the exploited individual may be her assigned partner; the individual that is let down is her contractual partner, and hence has voluntarily exposed herself to this risk. Behavioral effects are heterogeneous. About a quarter of participants from a standard student subject pool exhibit the hypothesized additional deterrent effect. This fraction is bigger than a third if participants interact with a random partner from somewhere in the world. (*JEL* B21, C91, D22, D43, K12, L12, L14)

1. Introduction

Most deep insights are simple—in retrospect. But gaining the insight takes ingenuity. And alerting others to the insight requires a sovereign

Helpful comments by the editor Albert Choi, two anonymous referees, Sebastian Kube and Pascal Langenbach on an earlier version are gratefully acknowledged.

American Law and Economics Review
https://doi.org/10.1093/aler/ahad007
Advance Access Publication on January 8, 2024

command of language. All of these hold for one of the terms coined by Nobel Prize winner Oliver Williamson. Entering a contractual relationship can have the effect of a "fundamental transformation" (Williamson, 1985). While ex ante multiple buyers can choose with whom of multiple sellers to trade, once the deal has been struck it transforms into a bilateral monopoly. The buyer must live with this one seller, and the seller must live with this one buyer. The effect obtains if, once the contract has been concluded, it would be technically impossible, prohibitively costly, or illegal, to turn to a different interaction partner.

One cause of this fundamental transformation, that Oliver Williamson has been particularly interested in, is transaction-specific investment (Williamson, 1985). Such investment is defined by being worthless, or at least much less valuable, if not used for fulfilling the contract in question, or further developing this business relationship. Yet this is only one potential cause of the transformation. It can also become impossible to turn away from a contracting partner if the first interaction required sharing sensitive information. A legal application stems from the prohibition, for a law firm, to represent the opponent of one of its clients. This legal rule effectively forces law firms and clients into stable relationships. The fundamental transformation can also be brought about strategically. A classic illustration is bundling, for instance of cheap copying machines with costly, proprietary toner cartridges (Adams and Yellen, 1976; McAfee, McMillan, and Whinston, 1989; Nalebuff, 2004). Finally, the fundamental transformation may result from an agreement between private parties. Take two neighbors. One family would wish to buy a larger family car, but cannot afford it. The other family has trouble taking the kids to school in the morning, as school and work hours do not match. The second family may offer to pay the first for taking care of the transportation to school. Anticipating the extra stream of income, the first can buy the more expensive car. But if the first stops taking the kids, the second must reorganize their professional lives. And if the second stops paying the first for taking the kids, the first has trouble repaying the car loan.

Once the fundamental transformation has taken place, both sides to the deal are "locked in" (Farrell and Shapiro, 1989; Liebowitz and Margolis, 1995; Farrell and Klemperer, 2007). They have made themselves vulnerable

to the fate and to the goodwill of their contracting partner. If their partner goes bankrupt or closes this line of business, they are in trouble. Trouble also looms large if a change in circumstances gives the contracting partner power to exploit the dependence. In the literature, this is called a holdup situation (Ellingsen and Johannesson, 2004; Hoppe and Schmitz, 2011; Dufwenberg, Smith, and Van Essen, 2013). A graphic illustration is a debate between Coase (2000) and a team most frequently represented by Benjamin Klein over the reason for General Motors to acquire Fisher Body in 1926 (Alchian and Demsetz, 1972; Klein, Crawford, and Alchian, 1978; Klein, 1988, 2000); Klein et al.argue that Fisher Body was taking advantage of GM's dependence on their bodies, which GM ended by vertical integration.

For an agent intended to maximize profit, lock-in is just a contingency. Such an agent compares the profit from staying alone with the expected benefit from entering the relationship. If the event space and probabilities are well defined, this is a simple exercise in calculation. If not, the agent replaces objective with subjective probabilities, and an objective with a subjective event space (Savage, 1954).

Yet from a behavioral perspective, the decision to enter the relationship may look more fraught. Understanding to which degree this is the case, and under which circumstances, is the topic of this paper. Does lock-in aversion deter efficient trade? There is quite a list of behavioral regularities that might have this effect. In this project, we use experimental methods to gauge the degree by which behavioral effects stand in the way of mutually beneficial, and hence efficient, trade.

Participants have the opportunity to profit from an exchange where the exchange involves a small risk of loss. They may choose to "sell" this opportunity and instead take a smaller but risk-free payout. We test each participant twice, on a *Baseline*, where an investment may fail, but there is no other participant involved, and additionally on one of four treatments. Choices in the first part have no material consequences for the second part. The first part of the experiment is the same for all participants.

Our control condition, *Baseline*, tests risk aversion alone. Participants decide how much they would pay to avoid a risk (because the risk is inherent to the fundamental transformation). Participants are then randomly

assigned to one of four additional conditions that each test additional reasons why people might want to avoid the fundamental transformation. The *Windfall* condition tests inequality aversion: the risk is still stochastic, but the result of a bad outcome for the subject is that another player randomly receives a windfall. The *Exploitation* condition gives room for let-down aversion: if the investment fails, the transfer is no longer automatic. It requires the decision of another, randomly determined participant to appropriate the first participant's loss. Each participant is potentially on the active side (may appropriate money from another participant) and on the passive side (another participant may appropriate money from her). But they are matched with a separate participant on either side. The *Bilateral* condition introduces a true holdup situation: two participants are matched ex ante. If the risk materializes for one of them, the respective other may decide to appropriate the loss (if the risk materializes for both of them, each may decide to appropriate the loss from the other). The fourth treatment (*Exchange*) implements a complete fundamental transformation. In this treatment, participants are also randomly matched ex ante. But participants are only exposed to risk if both of them have agreed to be bound by the contract. Otherwise, they receive their outside payment. If both agree and the risk materializes for one of them, the other may appropriate the loss.

Our measure for deterrence is the price at which a participant is willing to sell the opportunity to be in the contractual relationship. The further this price is below the expected value, the more participants shy away from mutually beneficial trade, for one of the behavioral reasons that we test. Since each participant decides twice (in the *Baseline*, where the risk is stochastic, and in one of the treatments), from each we have two prices. This not only allows us to take the idiosyncratic degree of risk aversion out of the equation; risk preferences are known to be heterogeneous (see only Holt and Laury, 2002). The difference between the price the participant requests in the *Baseline* and in the treatment provides us with a precise measure of additional behavioral effects when the risk results from interacting with another participant.

When deciding between a safe, but potentially less profitable, option and letting the fundamental transformation take place, the decision-maker

incurs a risk. In the *Baseline*, this risk is stochastic. No contractual partner is afraid of being taken advantage of. But they foresee that, in retrospect, entering the relationship may have been a bad idea. They could refrain from doing so because they are too strongly averse to risk. They might also construct the expected gains from trade as a reference point (Kahneman, 1992; Köszegi and Rabin, 2006; Hart and Moore, 2008), and anticipate loss aversion should the risk materialize (Kahneman, Knetsch, and Thaler, 1991; Tversky and Kahneman, 1991). Or they might be concerned that, should the risk materialize, they will regret having entered the regime. Regret might reduce their self-esteem, which they anticipate (Zeelenberg et al., 1996; Zeelenberg and Beattie, 1997; Van de Ven and Zeelenberg, 2011).

In a contractual relationship, one partner's loss can be another partner's gain. Even if this gain is not resulting from one partner strategically exploiting the other, they might still dislike that an exogenous event changes the intended balance of outcomes. Specifically, they might see this as an instance of disadvantageous inequity, in the sense of Fehr and Schmidt (1999), which they anticipate to dislike. This would explain a difference between the *Baseline* and the *Windfall* condition.

If the risk consists, at least partly, of becoming vulnerable to exploitation by the contracting partner, intentions come into play (Charness and Rabin, 2002). If a participant does not deem the contracting partner sufficiently trustworthy, she may loathe being let down and rather shy away from the relationship (Dufwenberg and Gneezy, 2000). This could explain a difference between the *Windfall* and the *Exploitation* conditions.

If two participants are assigned to each other ex ante, this might trigger the behavioral norm of reciprocity (Fehr, Gaechter, and Kirchsteiger, 1997; Fehr and Gächter, 2000; Perugini et al., 2003; Falk and Fischbacher, 2006). If the other participant appropriates the loss regardless, they may experience disutility from negative reciprocity (Brandts and Solà, 2001; Greco et al., 2019; Shaw, Barakzai, and Keysar, 2019). This could explain the difference between the *Exploitation* and *Bilateral* conditions.

Finally, if two participants have voluntarily entered the relationship, knowing about the risk of exploitation, and they are exploited, they may see this as a breach of trust (see only Berg, Dickhaut, and McCabe, 1995). They may loathe being betrayed (Bohnet and Zeckhauser, 2004; Bohnet

et al., 2008). This could explain the difference between the *Bilateral* and *Exchange* conditions.

Summing up, the experiment is designed to isolate the following four behavioral channels on which the fundamental transformation might deter trade: 1. another participant may gain a windfall profit if the risk materializes; 2. whether she does, depends on a choice she makes; she must thus decide to exploit the opportunity to her advantage; 3. the risk is reciprocal: the second participant who may potentially exploit the first is herself vulnerable to being exploited by the first participant; 4. both participants have voluntarily entered the relationship, knowing that it comes with the opportunity to exploit the partner and the risk of being exploited by her. Arguably Oliver Williamson's fundamental transformation is simultaneously characterized by all four effects, and by the behavioral effects they might trigger. In the *Windfall* treatment, there is only (1). In the *Exploitation* treatment, there are (1) and (2). In the *Bilateral* treatment, there are (1), (2), and (3). In the *Exchange* treatment, all four effects are present.

In our experiment, on average prices are slightly, but significantly below the expected value of the contract. For the *Baseline* and all treatments except *Windfall*, this is explained by risk aversion. In the *Windfall* condition, participants ask for a significantly higher price than in the *Baseline* in exchange for not being in the contractual relationship. Hence, the additional behavioral effect of inequality does not only not deter trade; participants on average even prefer this situation over one where no other participant gains if the risk materializes for themselves. This finding suggests that an alternative behavioral effect dominates: participants find it appealing that, if the risk materializes, money is not destroyed (does not go back to the experimenter), but is redistributed. In the *Exploitation* and *Bilateral* conditions, they ask for a significantly lower price than in the *Baseline*. In these treatments, we thus find the highest degree of deterrence. Note that this result can only be explained by behavioral effects. Unless participants expect each and every other participant to harm them when given the opportunity, the expected value of the contract is higher than in the *Baseline* (where harm is mechanical if the risk materializes). The comparison with the *Baseline* and the *Windfall* treatment suggests that participants dislike being the victim of intentional harm. Average prices in the *Exchange* condition are not significantly different from those in the *Baseline*.

Average effects do, however, mask pronounced heterogeneity. In all treatments, we find participants who ask for a lower price in the treatment than in the *Baseline* and others who ask for a higher price. Yet distributions differ. In *Windfall*, participants predominantly ask for more money. This fits a preference for another participant benefitting over money going back to the experimenter. The opposite effect dominates in the *Exploitation* and *Bilateral* treatments. In these treatments, intentions matter. Participants dislike exposing themselves to willful harm inflicted on them by another participant. Finally, in the *Exchange* treatment, the dominant reaction to treatment is the same as in the *Baseline*. In this treatment, apparently, behavioral effects pointing into opposing directions compete with each other: a participant risks being let down by a contractual partner, but she has voluntarily exposed herself to this risk and may trust her random counterpart to reciprocate. The result suggests that these effects cancel out.

We have run our original experiment with randomly selected, anonymously interacting student participants. Our choice of sample might have reduced external validity, as the typical student does not have much negotiation experience. Moreover, despite the precautions that are standard in economic experiments, participants might have seen themselves as a member of a community of students simultaneously present in the computer lab. To address both concerns, we have rerun the otherwise identical experiment on the online platform Prolific. On this platform, participants know that they are interacting with anonymous others from all over the world. We have constrained the sample to participants who are at least 30 years old, and who report to have negotiation experience. With this constraint, we did not only want to capture participants for whom "negotiation experience" was likely consequential. We, in particular, wanted to make it very unlikely that participants were students so that we could retest the effects with a truly different sample.

There are indeed clear differences between the original experiment and the second study on Prolific. Overall, participants on Prolific are less willing to accept any risk, be it stochastic or strategic. In the *Windfall* condition, participants no longer accept more risk than in the *Baseline*. There is again heterogeneity, but it is differently influenced by treatments. On Prolific, in the *Exploitation* and *Bilateral* conditions, more participants are willing to

increase their exposure to risk, compared with the *Windfall* condition; we had found the opposite in the original experiment.

A difference in correlations with post-experimental tests provides an explanation. In the original experiment, beliefs, and social value orientation were critical. This suggests that those participants who have increased their willingness to pay for playing the game have indeed seen the lab as a community of like-minded others who jointly aim at making the highest profit. If betrayal is excluded by design (i.e., in the *Windfall* condition), this holds for so many that we even find a significant overall effect. By contrast on Prolific, the scores from the test for loss aversion turn out critical. Those who are strongly averse to making a loss reduce their exposure to risk if this risk is strategic, that is, in the *Exploitation* and *Bilateral* treatments. Those who care little or not at all about making a loss even increase their willingness to pay for entering the game, compared with the *Baseline*. This suggests that, with greater social distance, participants focus on the possibility of exploitation. Some are not deterred by the fundamental transformation. They, to the contrary, see it as an opportunity for themselves to make a higher profit. Yet the more a participant dislikes being the victim, also in the Prolific sample the fundamental transformation deters otherwise mutually beneficial trade. We mean to stress the obvious: these are consistent explanations of the data, but with our data, we can only offer the explanation. We are not in a position to discriminate between this and alternative mental mechanisms.

Our evidence thus adds behavioral detail to Oliver Williamson's picture. The fundamental transformation is an obstacle to efficient trade even if, in expectation and when exclusively interested in profit, the ensuing risk of exploitation would be worth taking. The deterrent effect does not only result from risk aversion. It is compounded by aversion against being the victim of intentionally inflicted harm. Yet the additional, behavioral, deterrence is not universal. It only affects a fraction of the population. The size of this fraction depends on the degree of social distance. In the lab, the deterrent effect is most pronounced and affects about a third of the sample if the individual cannot signal her expectation that no exploitation will take place, that is, in the *Exploitation* and *Bilateral* conditions. In the *Exchange* condition, that is, if the participant must have voluntarily entered the contractual

relationship, the deterrent effect is no more pronounced than in the *Windfall* condition, where redistribution is imposed by the experimenter. This suggests that participants count on others in the room to reciprocate if they have knowingly made themselves vulnerable. By contrast on Prolific, more than a third of all participants reduce risk in the *Exchange* condition, while only less than a quarter do in the *Exploitation* and *Windfall* conditions. Apparently, with much greater social distance, there is little trust, but some appreciate the opportunity to exploit other participants.

One may wonder which of the two studies is more important from a governance perspective. For Oliver Williamson's model to apply, after the deal has been struck it must be costly or impossible for either party to break up the relationship. This can happen with wide social distance. If hardware has been subsidized, the provider may depend on the customer continuing to pay for maintenance, and the customer may depend on maintenance being reliable and affordable. But the classic situation that has motivated Oliver Williamson's research is characterized by a stable, long-term business relationship between partners who know each other very well. The relationship between a law firm and a client is typically also more personal, as is the relationship between the two neighbors in the other motivating example.

The remainder of the paper is organized as follows: in Section 2 we relate our experiment to the literature and define our contribution. In Section 3 we develop hypotheses. In Section 4 we report the details of the design. Section 5 is devoted to results from the original experiment. Section 6 reports the results from the second experiment on Prolific. Section 7 concludes with a discussion.

2. Contribution

There is a series of three related experimental papers that investigate Williamson's fundamental transformation. These papers start from the theoretical contribution by Hart and Moore (2008). The critical element of their model is mutual incompleteness of the contract. The buyer is free to choose a minimum price or to pay more. She can, however, exclude price adjustments by choosing a "rigid" contract. The seller either faces low or high production cost. She determines the buyer's profit by choosing

quality. The model is interested in "shading." At a small cost for herself, the seller can strongly reduce the buyer's profit. In behavioral terms, this is an instance of costly punishment (Güth, Schmittberger, and Schwarze, 1982; Andreoni, Harbaugh, and Vesterlund, 2003; Casari and Luini, 2009; Almenberg et al., 2011; Balliet, Mulder, and Van Lange, 2011). The model expects the propensity to shade to depend on the seller's reference point, which, in turn, is predicted to be determined by the choice of contract. The model assumes that the seller evaluates the price chosen by the buyer once the uncertainty about the state of nature has been resolved in the light of this reference point.

In the lab, the main prediction of the model is supported. There is little shading if the buyer chooses a "rigid contract." With this contract, there is no trade at all if the seller faces high production costs. If the buyer chooses a "flexible contract," in the bad state of Nature the buyer may adjust the price upwards. The less she does, the more the seller is inclined to produce low quality (Fehr, Hart, and Zehnder, 2011). Yet the difference in shading between contracts disappears if buyers and sellers are randomly matched, rather than determined by competition (Fehr, Zehnder, and Hart, 2009; Fehr, Hart, and Zehnder, 2011). By contrast, the predicted effects still go through if there is the additional possibility to conclude informal agreements or to renegotiate the contract after the state of nature is revealed (Fehr, Hart, and Zehnder, 2015).

These are important findings. But in a way, these experiments address the second question before the first. They investigate how contractual partners react once they find themselves in a bilateral monopoly. These papers only report in passing how the prospect of bilateral monopoly affects the choice of contract. If there is competition between buyers and sellers, about half of them choose the rigid contract, with 38% in the beginning, and 56% in the final period (Fehr, Hart, and Zehnder, 2011). If there is no competition, only 28% choose the rigid contract in the first experiment (Fehr, Hart, and Zehnder, 2011), and initially 21% in the second experiment. By the end of that experiment, this fraction goes down to 9% (Fehr, Zehnder, and Hart, 2009).[1] By contrast, we are chiefly interested in the choice of

1. In Fehr, Hart, and Zehnder (2015), the fraction of rigid contracts is not reported.

initial contract, and in isolating its behavioral determinants. We elicit participants' willingness to accept the risk of exploitation that is inherent in the fundamental transformation.

Further experimental papers are more remote. Blankenborg, Kaplan, and Miller (2012) find that the risk of a holdup situation only deters trade if the second mover has considerable bargaining power. Dufwenberg, Smith, and Van Essen (2013) implement a sequential game with three stages. At the first stage, the first mover can terminate the game, at the second the second mover can. If the second mover does, the outcome is efficient and both gain the same amount. The treatment difference is in the third stage. If the first mover has the option to punish the second mover, at a small cost to herself, she is considerably more likely not to terminate the game at the origin. Hoppe and Schmitz (2011) show that, against predictions from standard theory, option contracts can mitigate the holdup problem. Sloof, Sonnemans, and Oosterbeek (2004), Ellingsen and Johannesson (2004), and Erlei and Siemer (2014) are not interested in the formation of the bilateral relationship, but in investment choices in anticipation of later bargaining over the surplus.

Williamson has characterized the normative issue as the risk of holdup (Williamson, 1985). Experimental evidence has qualified the concern, showing that participants who have been taken advantage of are likely to strike back, even if this further reduces their payoff, and that this is at least partly anticipated (Dufwenberg, Smith, and Van Essen, 2013). Yet the power to strike back may be limited in the field and is ruled out by design in our experiment.

The behavioral literature on contracts (good overviews are Eisenberg, 2014; Koszegi, 2014; Zamir and Teichman, 2018) is not directly applicable. Oliver Williamson's concern is not being bound by a contract that, after the fact, turns out to be a bad deal. He is concerned about the contractual relationship. Once they have concluded the contract and adjusted to it, the parties depend on each other. Not because their opponent could legally enforce the contract, but because of transaction-specific investments that make it ex post irrational to leave the relationship, even if, anticipating this situation, it would have been better not to enter the relationship. This is also why contract remedies cannot work as a technology for sharing risk among the parties (cf. Polinsky, 1983).

In a more indirect manner, the following findings from the literature on contracts, however, hint at potential behavioral affects. Laypersons make a distinction between the unintentional and the intentional breach of contract. They deem it more blameworthy if breach of contract is motivated by the opportunity to sell the commodity at a higher price to a third party, rather than by an increase in cost (Wilkinson-Ryan and Baron, 2009). This supports the expectation that they particularly dislike if their contractual partner takes advantage of their specific investment.

There is also a link to the behavioral literature on efficient breach of contract. Contract may consistently be modeled as an option: either I fulfill the contract or I will pay expectation damages (Posner, 2009). Yet this is at odds with laypersons' moral intuitions (Lewinsohn-Zamir, 2012). In the same spirit, people who once entered a contractual relationship may feel let down if their partner exploits their vulnerability, resulting from entering the relationship.

The concern is also echoed by the behavioral literature on contracts that are specifically designed to exploit behavioral weaknesses of a contracting partner (summarized by Koszegi, 2014). One could argue that a partner letting down her counterpart in a contractual relationship exploits a "false belief" (Koszegi, 2014), namely the expectation that the partners will resolve contingencies in a spirit of mutual trust.

In their model, Herweg and Schmidt (2015) show why contracts become sticky if both parties are loss averse and the contract provides them with a reference point. This line of thought can be extended to a contractual relationship à la Williamson.

Further behavioral effects observed with contracts would only be relevant if they are anticipated. Relying on a finding by Hoffman and Wilkinson-Ryan (2013), Zamir and Teichman (2018) argue that there is a psychological shift between an ex ante perspective (dominated by cost-benefit analysis and a precautionary spirit) and an ex post perspective (characterized by a spirit of cooperation), also due to a tendency to reduce cognitive dissonance (I am in that relationship anyhow, so better put up with it). A deterrent effect could obtain if this effect is anticipated.

The effect could be exacerbated by a shift in the reference point (Köszegi and Rabin, 2006), and hence in either party's assessment of a legitimate

outcome, brought about by the contract (Hart and Moore, 2008). This effect has also been demonstrated experimentally (Fehr, Zehnder, and Hart, 2009; Fehr, Hart, and Zehnder, 2011), (also see Depoorter and Tontrup, 2012). If the parties anticipate the shift, they might be even more scared of entering the relationship.

3. Design

The main experiment has two parts. The first part is the same for all participants. They choose between a safe outside option and a risky investment. At this point, participants only know that the experiment has multiple parts, but not what these parts are about. The second part repeats the first. But now participants are randomly matched with one (*Bilateral, Exchange*) or two other participants (*Windfall, Exploitation*). The design thus combines a within-subjects design (first vs. second part) with a between-subjects design (one of the four treatments in the second part). Treatments are summarized in Table 1.

The *Baseline* and each of the treatments have two stages. In the first stage, participants choose between participating in the second (risky) stage and receiving a safe outside option. In the first stage, specifically, participants decide which amount they request for giving up the opportunity to participate in the second stage. We elicit their willingness to accept by way of a mechanism in the spirit of Becker, DeGroot, and Marschak (1964). A person participates in the second stage provided the minimum price she has chosen is the same or larger than the random price that the computer selects from a uniform distribution, in the range [0, 100]. Otherwise, she receives

Table 1: Treatments

	Human beneficiary	Room for exploitation	Room for reciprocal exploitation	Voluntary consent
Baseline				
Windfall	x			
Exploitation	x	x		
Bilateral	x	x	x	
Exchange	x	x	x	x

the number of experimental currency units (ECU) that the computer has randomly drawn, and this part of the experiment ends for her.[2] Provided she enters the second stage, she obtains 100 ECU with a probability of 80%, and 10 ECU with probability 20%. Hence the expected value of taking the lottery is 82 ECU.

In the second part of the experiment, participants are in one of four treatments. All treatments differ from the *Baseline* by the fact that, if the bad state of Nature obtains, another experimental participant gains a windfall. In the *Windfall* treatment, this effect is mechanical. By contrast, in the remaining treatments, in the bad state of Nature, another participant has the power to take 90 ECU. In the *Exploitation* treatment, the participant who has the power to take differs from a second participant from whom the first participant can take 90 ECU in case the bad state of Nature obtains (in this relationship). In the *Bilateral* and in the *Exchange* treatments, two experimental participants are matched. In the *Bilateral* treatment, matching is ex post. One other participant who has decided to accept the contract is randomly matched. By contrast in the *Exchange* treatment, matching is ex ante.[3]

In the interest of obtaining full data, we use the strategy method (Selten, 1967) if the payoff relevance in the bad state of Nature depends on the decision of another participant. We elicit this choice conditional on the opportunity presenting itself.

To preempt hedging and interference between choices in the first and the second part of the experiment, participants learn that either the first or the second part is paid out, with equal probability. All random draws are independent of each other. All are executed at the very end of the entire experiment. All feedback is withheld until participants have made all choices to preserve independence.

2. Please see the instructions in the Appendix for a more elaborate explanation of the mechanism.

3. If the second part is payoff relevant, the second stage is only implemented if both participants in the ex ante match have decided to participate in this stage. Otherwise, they receive the number that the computer has randomly chosen. This is also how we proceed if the number of participants in a session is odd for the one unmatched participant.

In the interest of having more scope for isolating the behavioral effects that motivate our hypotheses, after the main experiment, we run a series of standard tests: risk aversion, using the standard test by Holt and Laury (2002); loss aversion, using the test by Gächter, Johnson, and Herrmann (2007); regret aversion, using the test by Schwartz et al. (2002); social value orientation, using the standard test by Murphy and Ackermann (2014); trust and trustworthiness, adapting the test by Berg, Dickhaut, and McCabe (1995), and justice sensitivity (Schmitt et al., 2010). We finally ask which minimum price this participant believes one other participant to have chosen in the second stage of the experiment. In the *Exploitation*, *Bilateral*, and *Exchange* treatments, we also ask whether the participant believes that the other participant with whom they have been matched has decided to take 90 ECU from them should the opportunity present itself. Beliefs are incentivized. For detail, we refer to the instructions in the Appendix.

In a pilot, we had the impression that a substantial fraction of participants might not have understood the mechanism that we use to elicit their willingness to accept. We have reacted with a series of safeguards: 1. in the instructions we extensively explain the mechanism used to elicit willingness to accept; 2. we have participants solve two control questions; 3. we explicitly inform participants about the expected value of concluding the contract to avoid that results are influenced by calculation errors or the inability or unwillingness to calculate the profit-maximizing choice. We acknowledge that this number can work as an anchor. Yet if participants do the calculations, they have the same information. And most importantly, we are chiefly interested in the difference between choices in the *Baseline* and in the respective treatment. This dependent variable cannot be affected by a potential anchor, as the potential anchor is held constant in both parts of the experiment. For all details regarding the design of the experiment, we refer to the instructions in the Appendix.

The experiment has been run in the lab of the Max Planck Institute for Research on Collective Goods in Bonn. Participants have been invited with software hroot (Bock, Baetge, and Nicklisch, 2014). The experiment has been programmed with software zTree (Fischbacher, 2007).

Two hundred participants, most of whom had been students of various majors, participated who have been randomly selected from a pool of approximately 6,000 participants. One hundred and twelve (56%) were female. The mean age was 25.01 years. Participants on average earned 19.40€ (equivalent to 21.38$ on the first day of the experiment). As not all invited participants showed up, the number of participants varies per treatment. We have forty-seven participants in the *Windfall* treatment, fifty in *Exploitation*, fifty-four in *Bilateral*, and forty-nine in the *Exchange* treatment.

4. Hypotheses

In this part, we formulate the hypotheses that have motivated the design of the experiment. As hypotheses are only partly supported by the data, we discuss alternative behavioral channels when reporting results.[4]

4.1. Standard Preferences

If participants hold standard preferences themselves and expect other participants to hold standard preferences as well and to also expect that others expect them to hold such preferences, they expect that others will seize the opportunity to make a higher profit whenever it presents itself. Hence, if they interact with another human participant and the random draw makes it possible for this participant to exploit them, they expect this to happen. In the experiment, the probability is known with which exploitation is possible. Therefore, participants can calculate the expected value of the risky option. They are indifferent between this option and its certainty equivalent. This gives us

Hypothesis 1: a) Participants only prefer the safe outside option if its value exceeds the expected value of the risky option. b) There are no treatment differences.

4. As we have only preregistered the second study, when we had already seen the results from the first study, in our preregistration, we have more cautiously refrained from directed hypotheses, see https://osf.io/nksjq/?view_only=38b117e080d74e3c8fa896c9fa84e3f9.

4.2. Risk Aversion, Loss Aversion, and Regret Aversion

If participants are risk averse, their utility from a lottery is smaller than its expected value. On average, participants of experiments have been shown to be risk averse (Cartwright, 1971; Holt and Laury, 2002; Fullenkamp, Tenorio, and Battalio, 2003; Dohmen et al., 2011; Charness and Gneezy, 2012). This should affect their willingness to accept the fundamental transformation. In this perspective, avoiding the fundamental transformation can be interpreted as insuring oneself against being locked in. The effect should even be more pronounced if participants are, additionally, averse to losses, and if they interpret the sure outcome in case they avoid the fundamental transformation as a reference point (Tversky and Kahneman, 1991; Köszegi and Rabin, 2006). Participants might also anticipate that, should the risk materialize, they will regret having entered the relationship. Regret might reduce their self-esteem, which they anticipate (Zeelenberg et al., 1996; Zeelenberg and Beattie, 1997; Van de Ven and Zeelenberg, 2011).

These considerations yield

Hypothesis 2: a) The willingness to accept the fundamental transformation is below its expected value. b) There are no treatment differences.

4.3. Inequality Aversion

The *Windfall* treatment differs from the *Baseline* by the effect of the fundamental transformation. If the risk materializes, the participant who has been willing to lock herself in not only loses money herself. This money goes to another participant who, otherwise, would have had the same payoff as herself. It has been shown that many participants are averse to inequality, in particular, if it is to their detriment (Fehr and Schmidt, 1999; Blanco, Engelmann, and Normann, 2011).

Now with respect to a third participant, the decision-maker is herself in the position of potentially making a windfall gain. Yet those two relations are unrelated and the design keeps the probability that the risk materializes small. More importantly, the active participant has no influence on receiving a windfall gain, while it is in her power to prevent another participant from gaining extra money. Finally, if windfalls occur in both dimensions, the active participants potentially experience disadvantageous inequity in

one direction, and advantageous inequity in the other. Most participants are more sensitive to the former than to the latter. Hence even if both windfalls co-occur, the behavioral effects do not cancel out, but compound.[5]

We predict

Hypothesis 3: Willingness to accept the fundamental transformation is lower in the *Windfall* treatment than in the *Baseline*.

4.4. Let Down Aversion

In the *Windfall* treatment, if the risk materializes another participant gains extra money. Yet the resulting inequality is mechanical. This is different in the *Exploitation* treatment. In this treatment, whether the active participant loses money is a choice of another, randomly determined participant. The fundamental transformation thus puts the active participant at the mercy of another experimental participant. It has been shown that individuals are averse to being exploited (Dufwenberg and Gneezy, 2000).

We therefore predict

Hypothesis 4: Willingness to accept the fundamental transformation is lower in the *Exploitation* than in the *Windfall* treatment.

4.5. Holdup

In the *Exploitation* treatment, if the risk materializes, one participant is at the mercy of another. Yet the risk is unilateral. A may be exploited by B, but B does not run the risk to be exploited by A. A may exploit C, and D may exploit B, but these are unrelated contingencies. This is different in the *Bilateral* treatment. Nature matches two individuals who have (independently and with no information about the person with whom they will

5. Note that, with 90% probability, this result also holds if the participant in question does not compare payoff per relationship, but total payoff. If A takes from B, in this relation B earns 10. If she happens to have a chance to take from C, and does so, in that relationship, she earns 90. Her total earnings (from this part of the experiment) are thus 100. This is more than C, and less than A, unless C also has the opportunity to take and acts upon it, and A is the victim of taking from yet another participant. Either opportunity only presents itself with probability 10%. Both of them only present themselves simultaneously with probability 1%.

be matched) chosen to run the risk of the fundamental transformation. If the risk materializes for both of them, either of them may exploit the other. This design places participants in a holdup situation. The existing experimental evidence is tentative: under some circumstances trade with the prospect of a holdup becomes less likely (Ellingsen and Johannesson, 2004; Dufwenberg, Smith, and Van Essen, 2013). With our design, we have the possibility to discriminate between the risk of exploitation and the holdup situation and generate a cardinal measure for the difference between both.

Oliver Williamson's (informal) model transcends the mere risk of exploitation. He characterizes the situation as a "bilateral monopoly" (Williamson, 1985). The bilateral element might trigger the strong behavioral norm of reciprocity (Fehr, Gaechter, and Kirchsteiger, 1997; Fehr and Gächter, 2000; Perugini et al., 2003; Falk and Fischbacher, 2006) and of negative reciprocity in particular (Brandts and Solà, 2001; Greco et al., 2019; Shaw, Barakzai, and Keysar, 2019). When individuals perceive a violation of the norm of reciprocity, they feel the urge to strike back. If they anticipate this possibility, they might shy away from exposing themselves to the situation in the first place.

Based on these earlier findings, we expect

Hypothesis 5: Willingness to accept the fundamental transformation is lower in the *Bilateral* than in the *Exploitation* treatment.

4.6. Voluntary Trade

In Oliver Williamson's thinking, the fundamental transformation does not collapse with a holdup situation. He is not exclusively considering the situation ex post, where one individual is at the mercy of the other. Rather he is interested in the decision to bring the transformation about when it could have been avoided. It is not a move of Nature that creates the holdup situation; individuals bring it about by mutual consent. They decide to replace a situation where they remain independent with a bilateral monopoly. This is what he calls a "fundamental transformation" (Williamson, 1985). In the experiment, we capture this difference by a further condition. In the *Bilateral* treatment, we only match participants who have decided to run the risk of exploitation. Yet they have made this decision independently, and matching is at random. By contrast in the *Exchange*

treatment, participants are not only matched. They only enter the contractual regime if both of them agree to be bound. Behaviorally, the difference is one of trust. By accepting the fundamental transformation, the matched participants put trust into each other.

Participants might be even more deterred from entering the fundamental transformation, as the materialization of the risk is then also a breach of trust: they may loathe being betrayed (Bohnet and Zeckhauser, 2004; Bohnet et al., 2008). They may also be concerned that only participants who are planning to take self-select into the second stage of this part of the experiment. There is a substantial behavioral literature supporting this view. Betrayal aversion has been defined as the aversion to uncertainty resulting from the decision of another person, compared with stochastic risk (Bohnet and Zeckhauser, 2004; Humphrey and Mondorf, 2021). Empirically, it is pronounced (Koehler and Gershoff, 2003; Bohnet and Zeckhauser, 2004; Aimone and Houser, 2012). The effect obtains for male and female participants in Oman, Switzerland, Turkey, and the United States, but not in Brazil and China (Bohnet et al., 2008).

Provided these behavioral mechanisms are at work, we expect

Hypothesis 6: Willingness to accept the fundamental transformation is lower in the *Exchange* than in the *Bilateral* treatment.

5. Study 1: Results

5.1. Treatment Effects on Means

Figure 1 summarizes the treatment effects in the original experiment. In the *Baseline* and in all treatments, the average price participants demand for not entering the second stage, that is, the contract that is fraught with the risk of losing 90 ECU is slightly below its expected value of 82. The difference from the expected value is significant for the *Baseline* (*t*-test, $N = 200$, $P < 0.001$, one-sided) and for all treatments except *Windfall* (*Exploitation* $N = 50$, $P = 0.001$; *Bilateral* $N = 54$, $P = 0.0003$; *Exchange* $N = 49$, $P = 0.027$).[6] This rejects $\mathbf{H}_1$: on average, participants prefer to

6. We use a one-sided test since the alternative hypothesis $\mathbf{H}_2$ is directed, and expects price to be below its expected value. If we use a two-sided test, for the *Exchange* treatment, the difference to the expected value of 82 is only weakly significant, $P = .054$.

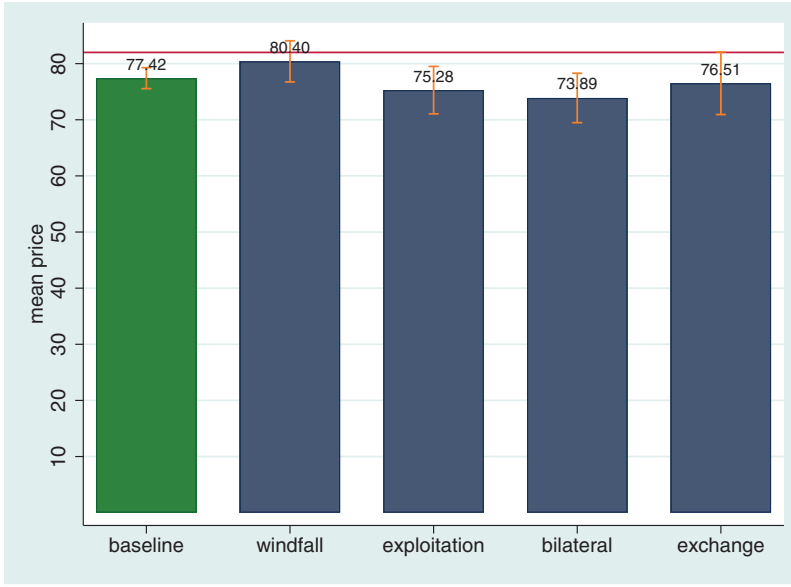

Figure 1. Treatment Effects on Prices Demanded for Not Entering the Contractual Relationship. Red Line at 82: Expected Value of the Contractual Relationship. Error Bars for 95% Confidence Interval.

exchange the contractual relationship against a certain payment that is slightly below the expected value of entering the relationship. Recall that we have explicitly told them the expected value and have explained what it means. Hence, in the *Baseline* and three of the treatments, participants knowingly accept a lower expected profit. This supports the competing hypothesis $\mathbf{H}_2$.

In $\mathbf{H}_3$, we had expected that choices in the *Windfall* treatment are motivated by inequality aversion and participants sell the opportunity to participate in the game at a price significantly below choices in the *Baseline*. We actually find the opposite. Participants on average ask for 3.830 ECU more, which is a significant difference (*t*-test testing the null hypothesis that this difference is zero, $N = 47$, $P = 0.0212$, two-sided).

In $\mathbf{H}_4$, we had expected that prices in the *Exploitation* treatment are significantly lower than in the *Windfall* treatment. We do indeed find a weakly significant effect ($P = 0.070$). As participants sell the opportunity to participate in the game at a lower price, this suggests that they loathe being

let down by another participant. In terms of the theory of social preferences, participants do care about intentions (Rabin, 1993) more than they care about a mere inequality of outcomes (Fehr and Schmidt, 1999). In the Appendix, we report balance with respect to all post-experimental tests across treatments. Between these two treatments, risk aversion and social value orientation differ at least at the 10% level. If we run regressions that control for these independent variables, either one by one or jointly, we replicate the weakly significant treatment effect. If we further take between-subjects heterogeneity out of the equation, by controlling for the choice this participant has made in the *Baseline*, and interact with this variable, we always find a treatment effect at the conventional level for significance ($P < 0.004$).[7]

In $\mathbf{H_5}$, we had expected that prices in the *Bilateral* treatment are lower than in the *Exploitation* treatment. This prediction is not supported by the data. We find no significant treatment difference, whether or not we control for or interact treatment with the choices in the *Baseline*. The nonresult does also not change if we control for the two variables that significantly differ between these two treatments, that is, risk aversion and regret aversion.[8] We do therefore not find an additional effect of aversion against a holdup situation.

In $\mathbf{H_6}$, we finally had predicted that prices in the *Exchange* treatment are lower than prices in the *Bilateral* treatment. We do, however, not find any significant difference, also not if we control for, or interact treatment with the choice the participant has made in the *Baseline*, or if we additionally control for the one post-experimental test that significantly differs between these two treatments, that is, social value orientation.[9]

The bottom line is that the fundamental transformation does deter trade. This is due to risk aversion. If someone else can benefit when a trading partner has bad luck, she likes that better than losing money to the experimenter. Yet if the fundamental transformation has occurred, benefit to someone else and exploitation coincide. Being the victim of exploitation is what participants dislike.

7. The code for these additional regressions is in the Supplementary Materials.
8. The code for all supplementary regressions is in the Supplementary Materials.
9. The code for all supplementary regressions is in the Supplementary Materials.

5.2. Heterogeneity

The previous analysis exclusively looks at central tendencies. Now from every participant, we have two choices: one in the *Baseline* and another in the respective treatment. This feature of the design makes it possible to dig deeper and analyze heterogeneity. Specifically, we compare, separately for each treatment, whether and how participants have changed choices between the *Baseline* and the respective treatment.

As Figure 2 shows, both the direction and the degree of heterogeneity differ markedly across treatments. In the *Windfall* treatment, 15% of all participants accept less risk in the treatment than in the *Baseline*, 40% accept more risk. By contrast, in the *Exploitation* and in the *Bilateral* treatments, 26% and 24%, respectively, accept more risk, whereas 38% and 37% accept less risk. Finally in the *Exchange* treatment 51% make the same choice as in the *Baseline*, and only 24% each shift toward more or less risk. The kernel density plot in the right panel adds cardinal information. If participants move into the direction of less risk, in the *Exchange* treatment, some of them move very far. In cardinal terms, there is also a descriptive difference between *Exploitation* and *Bilateral*. As the large peak in the neighborhood of zero shows, in *Exploitation*, participants either move clearly into the negative direction or they stick to their previous choice. By contrast, in the *Bilateral* treatment, there are more small negative shifts, which flattens the density curve.

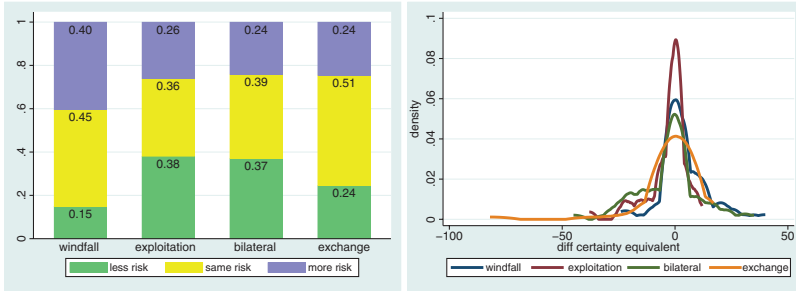

Figure 2. Heterogeneity of Reactions to Treatment. Left Panel: Fraction of Participants Who Choose Less, the Same or More Risk in the Treatment, Compared with the *Baseline*. Right Panel: Difference in Minimum Price Asked for Giving Up the Opportunity to Enter the Contractual Relationship, Treatment—*Baseline*, Kernel Density Plots.

5.3. Explanations

Our hypotheses are motivated by a series of behavioral effects. In this exploratory section, we exploit the richness of our data and, in particular, results from the post-experimental tests to find explanations for the observed effects.

We have found support for $\mathbf{H}_2$: participants sell the possibility to participate in the game below its expected value. The regression in Table 2 shows that this behavior is driven by risk aversion. The more a participant is averse to risk, the less she is likely in the *Baseline* to have a willingness to pay for participating in the game that is at or above its expected value. Note that the *Baseline* provides the cleanest test for this motivating force as, in the *Baseline*, social context does not have pecuniary consequences.

Against expectation, we have found that, in the *Windfall* treatment, participants have a significantly higher willingness to pay for participating in the game than in the *Baseline*. Table 3 shows that this effect is explained by social value orientation. The more the participant is willing to give up money for herself to reduce the payoff gap between herself and an anonymous counterpart, the more she prefers to play the game in the *Windfall* condition, compared with the *Baseline* where she faces a lottery with the same expected value. While this effect is only marginally significant ($P = 0.087$), we find an effect at the conventional significance level if we explain the difference in the willingness to pay between the *Windfall* treatment and the *Baseline* with a dummy that is 1 if the participant is willing to give up money for decreasing the payoff gap between herself and a passive outsider. This suggests that participants prefer money to go to another participant,

Table 2: Explaining Choices in the Baseline

Risk aversion	−0.187*
	(0.082)
Cons	0.536***
	(0.047)
N	200

Linear probability model. dv: dummy that is 1 if choice in the *Baseline* is 82 or higher. Risk aversion: score from Holt/Laury test. Inconsistent choices coded by frequency of choosing lottery with larger spread, as recommended by Holt/Laury. Standard errors in parenthesis.

* $P < 0.05$.

** $P < 0.01$.

*** $P < 0.001$.

Table 3: Explaining Difference Between Baseline and Windfall

	Model 1	Model 2
Social value orientation score	0.213^{+}	
	(0.122)	
Positive social value orientation		6.665*
		(3.109)
Cons	–0.025	0.143
	(2.705)	(2.313)
N	47	47

OLS. dv: WTP in *Windfall*—WTP in *Baseline*. Social value orientation score: angle in test by Murphy/Ackermann. Positive social value orientation: dummy that is 1 if the score is >7.15 (which is the score if participants refrain from inflicting harm on their counterparts in case they are indifferent themselves). standard errors in parenthesis.

*** $P < 0.001$.

** $P < 0.01$.

* $P < 0.05$.

$^{+}$ $P < 0.1$.

over money going back to the experimenter. This is related to an effect that, in the experimental literature, has sometimes been called a preference for efficiency (cf. Engelmann and Strobel, 2004)—which of course presupposes that welfare is defined narrowly, excluding the experimenter.

We finally aim at explaining the pronounced heterogeneity documented in Figure 2. We estimate a structural model, as one and the same participant simultaneously decides to increase or to decrease her exposure to risk when moving from the *Baseline* to the respective treatment. In line with the visual impression from Figure 2, treatments differ more pronouncedly in the propensity to reduce, rather than increase risk. In Table 4, we document a 22.9% higher willingness to reduce risk when a participant faces the prospect of another participant exploiting her, compared with the situation where a random participant gains a windfall. Participants are also 15.9% more likely to reduce risk in the *Bilateral* condition, but this effect is only weakly significant ($P = 0.079$). By contrast, the propensity to accept even more risk with social context is much less sensitive to treatment. The only weakly significant effect ($P = 0.076$) concerns the *Exchange* treatment. In this treatment, participants are 16.5% less likely to increase their exposure to risk when compared with the *Windfall* condition.

The choice between less, the same, or more risk in the treatment is most clearly influenced by participants' beliefs about the willingness to pay of others. The more they expect others to accept risk, the more they are inclined to

Table 4: Explaining Decision to Take Less or More Risk in Treatment, Compared with *Baseline*

	Less Risk	More risk
Exploitation	0.229**	−0.145
	(0.087)	(0.093)
Bilateral	0.159+	−0.151
	(0.091)	(0.096)
Exchange	0.023	−0.165+
	(0.088)	(0.093)
Belief	−0.007**	0.007**
	(0.002)	(0.003)
Risk aversion	−0.128+	0.076
	(0.075)	(0.079)
Loss aversion	−0.000	−0.010
	(0.020)	(0.022)
Regret aversion	−0.000	0.005
	(0.026)	(0.028)
Social value orientation	−0.006*	−0.001
	(0.003)	(0.003)
Trust	0.016+	−0.004
	(0.009)	(0.010)
Trustworthiness	−0.137	0.186
	(0.217)	(0.231)
Age	−0.008	0.005
	(0.007)	(0.007)
Male	−0.010	−0.073
	(0.064)	(0.068)
Cons	1.199***	−0.208
	(0.306)	(0.326)
N	185	

Linear Structural Model. less risk: dummy that is one if WTP in treatment < WTP in *Baseline*. more risk: dummy that is one if WTP in treatment > WTP in *Baseline*. Treatments: reference category *Windfall*. Belief: expected WTP of others in treatment. Risk aversion: score from Holt/Laury test. Inconsistent choices coded by frequency of choosing lottery with larger spread, as recommended by Holt/Laury. Loss aversion: inverse of the lottery at which participant switches to safe outcome of 0. Regret aversion: Schwartz Ward score. Social value orientation: angle in test by Murphy/Ackermann. Trust: choice as sender. Trustworthiness: (choice as a receiver when the sender sends 5 + choice as a receiver when the sender sends 10)/2. Standard errors in parenthesis.

*** $P < 0.001$.

** $P < 0.01$.

* $P < 0.05$.

+ $P < 0.1$.

expose themselves to more risk as well. The less they expect others to pay for entering the relationship, the more they reduce their own willingness to pay. Arguably, the more ambitious participants' beliefs, the more they see the

situation as a joint enterprise of themselves with one or two random partners. This explanation is consistent with the significant negative effect of social value orientation on the decision to reduce risk. The more a participant cares about the relative well-being of others, the less she reacts to the introduction of social context with reducing her exposure to (exploitation) risk.

Two further weakly significant effects are less straightforward. The more a participant is averse to (financial) risk, the less she is likely to reduce her exposure to risk in the treatment ($P = 0.088$). This could imply that the aversion to financial risk becomes less important if this risk originates in a social context and consists of another participant getting extra money. The more a participant is trusting, the more she is likely to reduce risk in the treatment ($P = 0.075$). This statistical effect likely is an artifact of the highly significant positive correlation between social value orientation and trust ($r = 0.393$, $P < 0.001$). Hence both explanatory variables likely pick up the same behavioral inclination.

6. Study 2: Increasing Social Distance

6.1. Design

Study 1 has been run in the computer lab of the Bonn Max Planck Institute, with participants from the large pool of more than 6,000 registered participants that the institute shares with the university. Most participants have been students.[10] Participants knew about this composition of the subject pool. One could, therefore, be concerned that the results do not carry over to interactions with greater social distance. Moreover, external validity would be greater if one tests participants with negotiation experience.

In the interest of addressing these concerns, we have rerun the (otherwise identical) experiment on the online platform Prolific. We have asked the platform to only invite participants who are at least 30 years old, are fluent in English, and report to have negotiation experience. Prolific reports that this holds for 10,667 participants. Actually, 94.42% of our participants indicate that they have negotiated their employment relation, 72.56% a

10. Fourteen of 200, that is, 7% have indicated that they are not enrolled at the university.

rental agreement, 57.67% a bank loan, 57.21% a business deal, 43.72% buying or selling a house, and 33.49% a car deal.

We have also translated the experiment into English, which has allowed us to reach a participant pool from all over the world. Prolific participants expect to meet other participants over large distances. Actually, we had 92 participants from the United Kingdom, 13 from the United States, and 110 from other countries. Participants were on average 39.08 years old (min thirty and max sixty-eight). Seventy of them were female, 144 male and one has indicated gender as diverse. We had fifty-seven participants in the *Windfall* condition, fifty-four in the *Exploitation* condition, fifty-two in the *Bilateral* condition, and fifty-four in the *Exchange* condition. Participants on average earned 12.96 £ (equivalent to 17.83 $ on the first day of the experiment). We have preregistered this second study. As in Study 1, most of our directed hypotheses have not been supported by the data, we have preregistered the more cautious undirected hypothesis that there are differences between treatments.[11]

6.2. Treatment Effects on Means

Figure 3 summarizes choices. As in the original experiment, descriptive differences between treatments in prices demanded for not entering the contractual relationship are small. Comparing with Figure 1, two differences between the original experiment and the second study are apparent: prices are further below the theoretical optimum when assuming risk neutrality. Willingness to enter the relationship is not increasing in the *Windfall* condition, compared with the *Baseline*. Choices in the *Baseline* are significantly below the theoretical benchmark of eighty-two (that we had made explicit in the instructions), *t*-test $N = 217$, $P < 0.001$. This supports $\mathbf{H_2}$. We also find weakly significant support for $\mathbf{H_6}$: in the *Exchange* condition, participants sell the opportunity to participate in the contractual relationship at a price below the one they require in the *Bilateral* condition (*t*-test $N = 104$, $P = 0.0599$, one-sided).[12] We do not find support for the remaining hypotheses.

11. https://osf.io/nksjq/?view_only=38b117e080d74e3c8fa896c9fa84e3f9, preregistration on September 1, 2021, prior to the first session of this second study.

12. If we use a two-sided test, which would be in line with the more cautious undirected formulation of the hypotheses in the preregistration, this effect is insignificant as well.

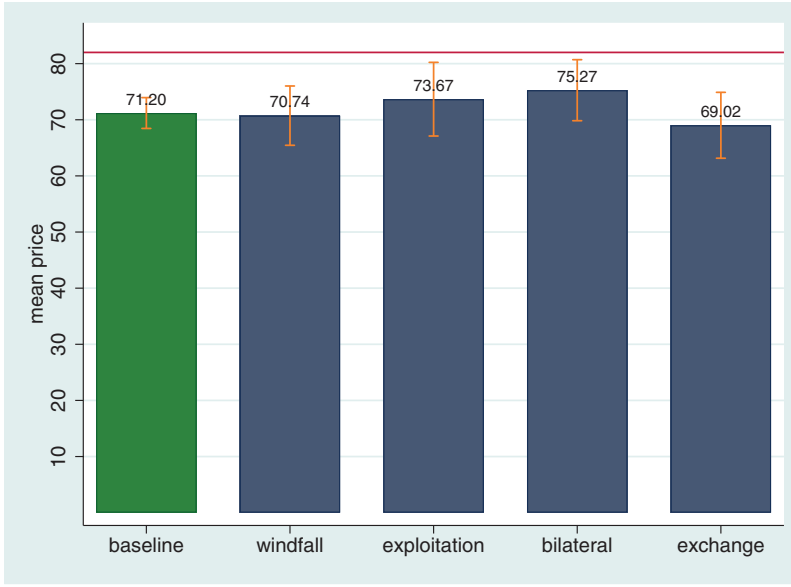

Figure 3. Treatment Effects on Prices Demanded for Not Entering the Contractual Relationship: Study 2. Red Line at 82: Expected Value of the Contractual Relationship.

6.3. Heterogeneity

Given the data from the original experiment, we expect pronounced heterogeneity in the reactions to treatment. This is indeed what we find, Figure 4. In the second study, many participants reduce their exposure to risk when moving from the *Baseline* to the *Windfall* treatment, and many increase their exposure to risk when moving from the *Baseline* to the *Exploitation* or the *Bilateral* condition.

6.4. Differences Between Samples

The second study is not a strict replication of the first. Rather with the help of a second experiment, we check whether the results from a study with (mostly) university students from the same university extrapolate to a setting where everything participants know is that they interact with others somewhere in the world who have agreed to earn a bit of money by

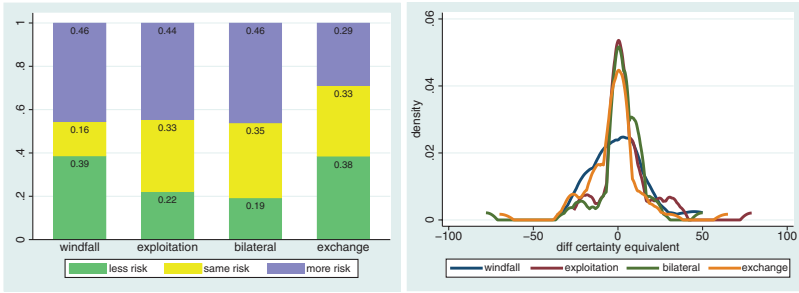

Figure 4. Heterogeneity of Reactions to Treatment in Study 2. Left Panel: Fraction of Participants Who Choose Less, the Same or More Risk in the Treatment, Compared with the *Baseline*. Right Panel: Difference in Minimum Price Asked for Giving up the Opportunity to Enter the Contractual Relationship, Treatment—*Baseline*, Kernel Density Plots.

being experimental subjects. All experiments, by their very design, are only analogous to the real-world situation that one wants to understand. This is the price one has to pay for identification through random assignment to treatment. The second study investigates generalizability in one dimension, by switching to a radically different subject pool: participants come from anywhere in the world, have no personal information whatsoever about the persons with whom they interact, are older, and report having negotiation experience. As Figure 5 suggests, the difference in samples does indeed matter. Differences are pronounced. As we have held the design of the experiment constant, and as we have randomly selected participants from much larger pools at the University of Bonn on the one hand, and from the even larger pool of participants who have agreed to be experimental subjects on Prolific, we can tentatively infer that the difference in the composition of the subject pool has been causal for the differences in outcomes. In the following, we offer a much larger social distance as a consistent explanation for the difference in outcomes. But this is of course an ex post rationalization. As very often in experimental research, one finding provides the research question for the next: which are the determinants of (perceived) social distance, and which are its effects on social interaction? This must be left to future research.

In which ways do results differ between platforms? As Figure 5 shows, in the lab, very few participants (14.89%) reduce their exposure to risk

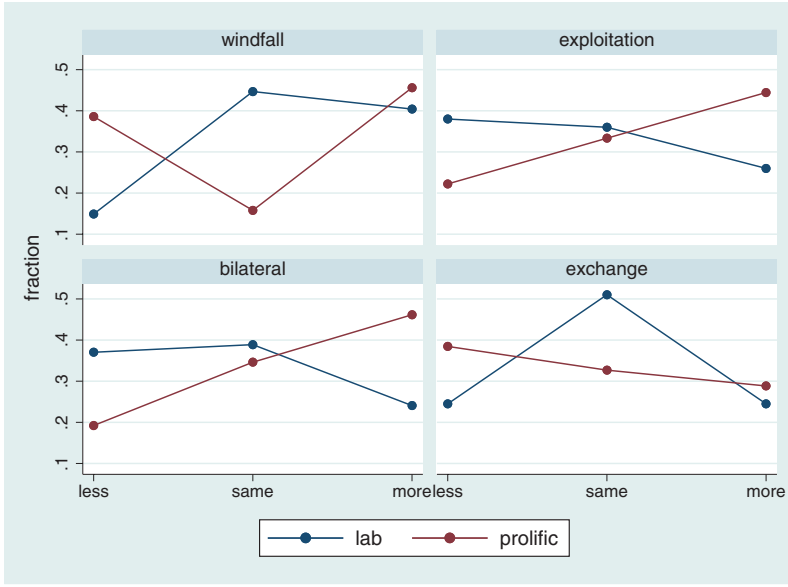

Figure 5. Comparing Reactions to Treatment between Lab and Prolific. Fraction of Participants, in the Respective Condition, Who Either Reduce ("less"), Increase ("more") or Do Not Change Their Exposure to Risk, Compared With the *Baseline*.

(ask for a lower price for selling the opportunity to participate in Stage 2) when moving from the *Baseline* to the *Windfall* treatment. The largest fraction (44.68%) sets the same price, closely followed by those who do even ask for more money in the *Windfall* condition (40.43%). On Prolific, almost all participants react to treatment; the fraction of participants who set the same price as in the *Baseline* is very small (15.79%). Participants essentially split into two groups. The slightly larger group (45.61%) are more interested in participating in the game (ask for a higher price to sell the opportunity to participate) than in the *Baseline*. But a not much smaller group (38.60%) is less willing to expose themselves to risk if this gives another anonymous participant a windfall profit.

In neither sample do we observe much of a difference between the *Exploitation* and the *Bilateral* conditions. Obviously, participants do not find it important whether the person who might exploit them is the same person whom they might exploit, should the opportunity present itself. But we find pronounced differences between the two samples. In the lab,

participants do not find these two conditions appealing. They are least likely to increase their exposure to risk, compared with the *Baseline* (24.07% do so in the *Bilateral* condition, and 26% in the *Exploitation* condition). The remaining participants split about equally between setting the same price as in the *Baseline*, and even reducing the price, and thereby their exposure to risk (36% and 38% in the *Exploitation* condition, and 38.89% and 37.04% in the *Bilateral* condition). The picture is very different on Prolific. Few Prolific participants reduce their exposure to risk, compared with the *Baseline*, when in the *Exploitation* (22.22%) or *Bilateral* condition (19.23%). A larger fraction asks for the same price as in the *Baseline* (33.33% in the *Exploitation* condition, 34.62% in the *Bilateral* condition). Yet in both treatments, the largest fraction finds the risk that another participant appropriates money more appealing than the risk that the money goes back to the experimenter (44.44% in the *Exploitation* condition and 46.15% in the *Bilateral* condition).

We also find a pronounced difference between samples when comparing choices in the *Baseline* with choices in the *Exchange* condition. In the lab, the majority (51.02%) eventually set the same price as in the *Baseline.* Equally sized minorities reduce and increase their exposure to risk (24.49% each). The pattern is very different on Prolific. Here, most participants (38.46%) set a lower price than in the *Baseline*, and hence increase the probability of not participating in the second stage of the experiment. A slightly smaller fraction (32.69%) ask for the same price as in the *Baseline*. The fraction of participants who are more eager to play the game and hence set a higher price (28.85%) is the smallest.

6.5. Explaining the Differences Between Samples

It is revealing to compare the structural models in Tables 4 and 5. We find support for the descriptive observation that the original experiment and the replication have induced pronouncedly different patterns. In the original experiment, participants are significantly more likely to reduce their exposure to risk when facing the risk of exploitation, compared with the *Windfall* condition where the extra gain for another participant results from a random draw. By contrast, in the second study, participants are less

likely to reduce their risk exposure in the *Exploitation* condition, compared with the *Windfall* condition. By contrast, the effect of the *Exchange* condition is consistent across studies: both in the original experiment and in the replication, participants are significantly less likely to increase their exposure to risk in the *Exchange* condition, compared with the *Windfall* condition.

Comparing the effect of explanatory variables hints at the source of the difference between the two studies. In the original experiment, beliefs and social value orientation had strong effects. By contrast, in the second study, the only explanation that is significant at conventional levels is the score from the test for loss aversion. If a participant scores high on this test (is pronouncedly averse to incurring a loss), she is substantially more likely to accept less risk and substantially less likely to accept more risk in the treatment, rather than the *Baseline*. In the test for loss aversion, the participant faces a choice between the safe outcome of 0 and a lottery with a 50% chance to gain 60, or to lose an amount between 20 and 70 ECU. The score is 0 if a participant even prefers the lottery if it has a negative expected value (she may lose 70). Hence, the lower the score, the more the participant cherishes the opportunity to make a positive gain. The significant effect of the score may, therefore, also be interpreted as an indicator of opportunity seeking. The more a participant is interested in not missing the opportunity for a higher profit, the more she increases her exposure to risk in the treatment, compared with the *Baseline*; recall that the opportunity to gain an extra 90 ECU is how all treatments differ from the *Baseline.*

Comparing Tables 4 and 5, therefore, suggests that social distance is indeed critical. In the student pool of the same university, the sense of solidarity prevails. Even if the recipient remains anonymous, participants prefer a situation in which, with some small probability, what one loses oneself helps another participant. By contrast, when interacting with a random stranger somewhere in the world, the opportunity to make an extra profit is critical, even if it comes at the risk of being exploited.

Yet if, on Prolific, participants predominantly care about the opportunity to make some extra money by exploiting another participant, why does this effect not play itself out in the *Exchange* condition? Now recall the

Table 5: Explaining Decision to Take Less or More Risk in Treatment, Compared with *Baseline*: Study 2

	Less risk	More risk
Exploitation	−0.171*	0.000
	(0.087)	(0.092)
Bilateral	−0.144	−0.078
	(0.089)	(0.094)
Exchange	−0.007	−0.203*
	(0.088)	(0.094)
Belief	−0.003+	−0.001
	(0.0015)	(0.002)
Risk aversion	−0.005	−0.075
	(0.058)	(0.062)
Loss aversion	0.041*	−0.055*
	(0.021)	(0.022)
Regret aversion	0.003	0.043
	(0.026)	(0.028)
Social value orientation	0.012	−0.173
	(0.129)	(0.137)
Trust	0.008	−0.012
	(0.011)	(0.012)
Trustworthiness	−0.233	0.300
	(0.203)	(0.217)
Age	−0.000	0.002
	(0.004)	(0.004)
Male	0.002	−0.087
	(0.068)	(0.072)
Cons	0.446	0.662*
	(0.312)	(0.333)
N	204	

Linear Structural Model. Less risk: dummy that is one if WTP in treatment < WTP in *Baseline*. More risk: dummy that is one if WTP in treatment > WTP in *Baseline*. Treatments: reference category *Windfall*. Belief: expected WTP of others in treatment. Risk aversion: score from Holt/Laury test. Inconsistent choices coded by frequency of choosing lottery with larger spread, as recommended by Holt/Laury. Loss aversion: inverse of lottery at which participant switches to safe outcome of 0. Regret aversion: Schwartz Ward score. Social value orientation: angle in test by Murphy/Ackermann. Trust: choice as sender. Trustworthiness: (choice as receiver when sender sends 5 + choice as receiver when sender sends 10)/2. Standard errors in parenthesis.

* $P < 0.05$.

** $P < 0.01$.

*** $P < 0.001$.

+ $P < 0.1$.

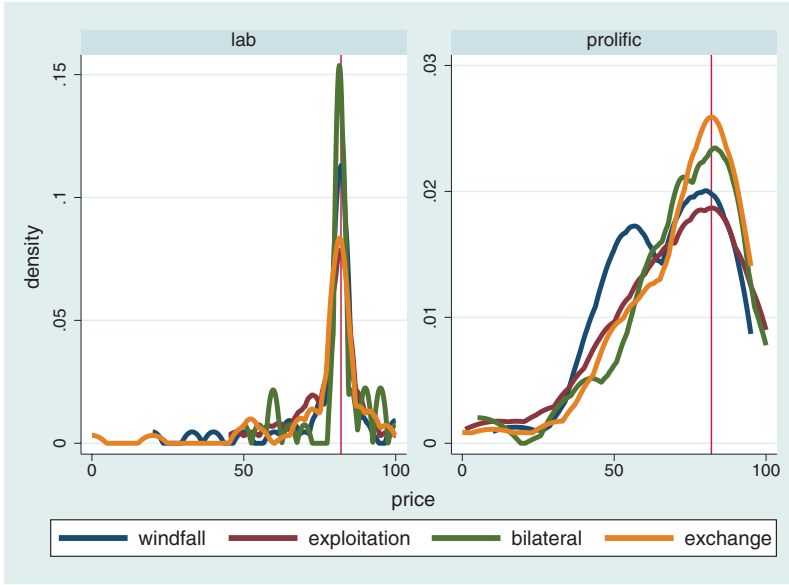

Figure 6. Beliefs About Prices Asked by Other Participants for Selling the Opportunity to Participate in Stage 2.

difference between *Exploitation* or *Bilateral* conditions on the one hand, and the *Exchange* condition on the other hand: in the latter condition, the second stage of the game is only played out if both participants of the previously randomly matched group of two have agreed to be bound by the contract. Figure 6 shows that, in all treatments, Prolific participants expect the remaining participants in the experiment to set much lower prices than in the lab. Recall that, irrespective of the sample, we have explicitly informed participants that they maximize expected profit when setting a price of 82. In the lab, a very large majority of participants expected their fellow participants to follow this advice. By contrast, Prolific participants have been much more skeptical. Now in other treatments (most pronouncedly in the *Windfall* condition) they have been even more skeptical than in the *Exchange* condition. Yet the only condition in which this belief directly matters for their own choice is the *Exchange* condition. If their designated counterpart has preferred to take the safe outside option, the participant in question is deprived of the opportunity for exploitation.

7. Discussion

Oliver Williamson has introduced the fundamental transformation as an incentive problem and as a challenge for the governance of the relationship. In this paper, we use a lab experiment to investigate whether the ensuing challenge for efficient trade is aggravated by behavioral effects. On average, the additional behavioral effects are not strong. The risk of exploitation is a risk with a pecuniary value. In the original lab experiment, participants mildly (but significantly) dislike this risk. The fact that they sell the opportunity inherent in the fundamental transformation at a price below its expected value indicates that participants are sensitive to this risk. If we only consider averages, the only additional behavioral effect obtains in the *Windfall* condition. In this condition, participants accept even more risk than in the *Baseline*, where the risk is stochastic. This suggests that they prefer if another participant makes some extra money on their account, over money going back to the experimenter. Yet this effect disappears as soon as the risk consists of the possibility that the matched participant decides to take money from the participant who has exposed herself to this risk. This suggests that being generous with others on the one hand, and aversion to being at the mercy of another person, cancel out. If the imbalance in outcomes results from an intentional act by another person, participants no longer want to be generous.

Yet we see a more pronounced deterrent effect once we allow for reactions to treatment to be heterogeneous. In all treatments, a fraction of participants accept less risk in the treatment than in the *Baseline*. These participants have been deterred by behavioral effects. Participants shifting into the direction of accepting less risk exist in the *Windfall* condition, yet they are much more frequent, and the size of the effect is more pronounced, in the *Exploitation* and *Bilateral* conditions. This indicates that a substantial minority of our participants would rather forego the opportunity of a higher profit than exposing themselves to intentional exploitation. In the *Exchange* condition, the fraction of participants exhibiting behavioral deterrence is again smaller, though. This suggests that trust can at least partly mitigate the additional deterrent effect. The story motivating this paper thus has to be qualified. The prospect of the fundamental transformation does deter trade, largely for the behavioral reasons that we hypothesized. Yet this only

holds for a fraction of individuals. In the original experiment, this fraction is about a quarter of the participants exposed to a stylized fundamental transformation.

The second study on Prolific provides a further qualification. Even if we maximize social distance, by matching participants with others anywhere in the world, a substantial fraction of them increases their exposure to risk if this risk is strategic, rather than stochastic. Yet in the original experiment, choices are explained by beliefs and social value orientation. This suggests a community interpretation. Many participants in the student lab seem to see themselves as part of a session-wide joint venture. By contrast, in the second study, only loss aversion explains heterogeneity. Participants with a high loss aversion score move away from risk if it is strategic. By contrast, those low in loss aversion move toward risk if it is strategic. They pay for the opportunity to exploit others. With greater social distance, trust and solidarity are not expected. Participants are attracted by the opportunity to exploit others or they are deterred by the prospect of being the victim. Arguably, in the *Exchange* condition, the same effect does not play itself out as participants are too skeptical about others participating in the interaction and hence about the opportunity to exploit another participant.

One may wonder which of the two studies is closer to the normative problem to which Oliver Williamson has alerted scholars and contract designers. Fundamental transformations exist over large social distance. Take a franchise contract. The firm that establishes the brand name may have little personal contact with a multitude of franchisees. If the franchisor is careless, the franchisee cannot do much to protect her business. On the other hand, the individual franchisee can cheat on quality and put the entire brand name at risk. Arguably if the relationship remains that distant, there is not much room for the development of mutual trust. But many of the illustrations discussed in the introduction refer to a situation where both sides of the relationship know each other very well. The lab data suggests that, if this is the case, additional behavioral effects deter at least a substantial fraction of individuals from entering the relationship in the first place.

In the interest of isolating behavioral channels, we have simplified the situation. In the field, the fundamental transformation is typically associated with a longer-term relationship. The prospect of being tied to another

person for a longer time might have an additional deterrent effect. But it, on the other hand, might also make it easier to accept the transformation, as the risk is not immediate. There might also be illusion of control (Langer, 1975; Presson and Benassi, 1996) or a better-than-average effect (Larrick, Burson, and Soll, 2007; Brown, 2012). Both effects might make individuals more inclined to accept the fundamental transformation if they had a true choice of contracting partners, and have picked theirs voluntarily. Either effect might make the individual believe that she can handle the risk. A further difference is the absence of social context. In the field, contracting partners are likely to collect information about each other before entering the relationship. While the relationship is ongoing, they are likely to monitor each other and to try to react to warning signals about future exploitation. They might also try to muster threat power to preempt exploitative moves. For all of these reasons, social context might make it easier for individuals to enter a potentially dangerous contractual relationship.

We do, therefore, not see our results as definitive. But our data show that the effects of the fundamental transformation transcend incentive effects. Legal and social rules meant to contain the risk of exploitation have the additional benefit of reducing the incidence of behavioral deterrence, resulting from the risk of exploitation.

Supplementary material

Supplementary material is available at *American Law And Economics Review* Journal online.

References

Adams, William James, and Janet L. Yellen. 1976. "Commodity Bundling and the Burden of Monopoly," 90 *Quarterly Journal of Economics* 475–98.

Aimone, Jason A., and Daniel Houser. 2012. "What You Don't Know Won't Hurt You. A Laboratory Analysis of Betrayal Aversion," 15 *Experimental Economics* 571–88.

Alchian, Armen A., and Harold Demsetz. 1972. "Production, Information Costs, and Economic Organization," 62 *American Economic Review* 777–95.

Almenberg, Johan, Anna Dreber, Coren L. Apicella, and David Rand. 2011. "Third Party Reward and Punishment. Group Size, Efficiency and Public Goods," in

Nicholas Palmetti and Jennifer P. Russo, eds., *Psychology of Punishment*. New York, NY: Nova.

Andreoni, James, William T. Harbaugh, and Lise Vesterlund. 2003. "The Carrot or the Stick: Rewards, Punishments and Cooperation," 93 *American Economic Review* 893–902.

Balliet, Daniel, Laetitia B. Mulder, and Paul A.M. Van Lange. 2011. "Reward, Punishment, and Cooperation: A Meta-analysis," 137 *Psychological Bulletin* 594–615.

Becker, Gordon M., Morris H. DeGroot, and Jacob Marschak. 1964. "Measuring Utility by a Single-Response Sequential Method," 9 *Behavioral Science* 226–32.

Berg, Joyce, John Dickhaut, and Kevin McCabe. 1995. "Trust, Reciprocity, and Social History," 10 *Games and Economic Behavior* 122–42.

Blanco, Mariana, Dirk Engelmann, and Hans-Theo Normann. 2011. "A Within-Subject Analysis of Other-Regarding Preferences," 72 *Games and Economic Behavior* 321–38.

Blankenborg, Dieter, Todd Kaplan, and Timothy Miller. 2012. "A Simple Economic Teaching Experiment on the Hold-Up Problem," 43 *Journal of Economic Education* 377–85.

Bock, Olaf, Ingmar Baetge, and Andreas Nicklisch. 2014. "hroot: Hamburg Registration and Organization Online Tool," 71 *European Economic Review* 117–20.

Bohnet, Iris, Fiona Greig, Benedikt Herrmann, and Richard Zeckhauser. 2008. "Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States," 98 *American Economic Review* 294–310.

Bohnet, Iris, and Richard Zeckhauser. 2004. "Trust, Risk and Betrayal," 55 *Journal of Economic Behavior & Organization* 467–84.

Brandts, Jordi, and Carles Solà. 2001. "Reference Points and Negative Reciprocity in Simple Sequential Games," 36 *Games and Economic Behavior* 138–57.

Brown, Jonathon D. 2012. "Understanding the Better than Average Effect: Motives (Still) Matter," 38 *Personality and Social Psychology Bulletin* 209–19.

Cartwright, Dorwin. 1971. "Risk Taking by Individuals and Groups. An Assessment of Research Employing Choice Dilemmas," 20 *Journal of Personality and Social Psychology* 361–78.

Casari, Marco, and Luigi Luini. 2009. "Cooperation Under Alternative Punishment Institutions: An Experiment," 71 *Journal of Economic Behavior & Organization* 273–82.

Charness, Gary, and Uri Gneezy. 2012. "Strong Evidence for Gender Differences in Risk Taking," 83 *Journal of Economic Behavior & Organization* 50–8.

Charness, Gary, and Matthew Rabin. 2002. "Understanding Social Preferences with Simple Tests," 117 *Quarterly Journal of Economics* 817–69.

Coase, Ronald H. 2000. "The Acquisition of Fisher Body by General Motors," 43 *Journal of Law and Economics* 15–32.

Depoorter, Ben, and Stephan Tontrup. 2012. "How Law Frames Moral Intuitions. The Expressive Effect of Specific Performance," 54 *Arizona Law Review* 673–717.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences," 9 *Journal of the European Economic Association* 522–50.

Dufwenberg, Martin, and Uri Gneezy. 2000. "Measuring Beliefs in an Experimental Lost Wallet Game," 30 *Games and Economic Behavior* 163–82.

Dufwenberg, Martin, Alec Smith, and Matt Van Essen. 2013. "Hold-Up: With a Vengeance," 51 *Economic Inquiry* 896–908.

Eisenberg, Melvin A. 2014. "Behavioral Economics and Contract Law," in Eyal Zamir and Doron Teichman, eds., *The Oxford Handbook of Behavioral Economics and the Law*. Oxford: Oxford University Press.

Ellingsen, Tore, and Magnus Johannesson. 2004. "Is There a Hold-Up Problem?," 106 *Scandinavian Journal of Economics* 475–94.

Engelmann, Dirk, and Martin Strobel. 2004. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments," 94 *American Economic Review* 857–69.

Erlei, Mathias, and J. Philipp Siemer. 2014. "Endogenous Property Rights in a Hold-Up Experiment," 65 *Metroeconomica* 237–70.

Falk, Armin, and Urs Fischbacher. 2006. "A Theory of Reciprocity," 54 *Games and Economic Behavior* 293–315.

Farrell, Joseph, and Paul Klemperer. 2007. "Coordination and Lock-In: Competition with Switching Costs and Network Effects," in Mark Armstrong and Robert H. Porter, eds., *Handbook of Industrial Organization*. Amsterdam, North-Holland: Elsevier.

Farrell, Joseph, and Carl Shapiro. 1989. "Optimal Contracts with Lock-in," 79 *American Economic Review* 51–68.

Fehr, Ernst, and Simon Gächter. 2000. "Fairness and Retaliation: The Economics of Reciprocity," 14 *Journal of Economic Perspectives* 159–82.

Fehr, Ernst, Simon Gaechter, and Georg Kirchsteiger. 1997. "Reciprocity as a Contract Enforcement Device: Experimental Evidence," 65 *Econometrica* 833–60.

Fehr, Ernst, Oliver Hart, and Christian Zehnder. 2011. "Contracts as Reference Points. Experimental Evidence," 101 *American Economic Review* 493–525.

Fehr, Ernst, Oliver Hart, and Christian Zehnder. 2015. "How Do Informal Agreements and Revision Shape Contractual Reference Points?," 13 *Journal of the European Economic Association* 1–28.

Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation," 114 *Quarterly Journal of Economics* 817–68.

Fehr, Ernst, Christian Zehnder, and Oliver Hart. 2009. "Contracts, Reference Points, and Competition: Behavioral Effects of the Fundamental Tansformation," 7 *Journal of the European Economic Association* 561–72.

Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-Made Economic Experiments," 10 *Experimental Economics* 171–8.

Fullenkamp, Connel, Rafael Tenorio, and Robert Battalio. 2003. "Assessing Individual Risk Attitudes Using Field Data From Lottery Games," 85 *Review of Economics and Statistics* 218–26.

Gächter, Simon, Eric J. Johnson, and Andreas Herrmann. 2007. "Individual-Level Loss Aversion in Riskless and Risky Choices". Mimeo.

Greco, Lindsey M., Jennifer A. Whitson, Ernest H. O'Boyle, Cynthia S. Wang, Joongseo Kim. 2019. "An Eye for an Eye? A Meta-Analysis of Negative Reciprocity in Organizations," 104 *Journal of Applied Psychology* 1117–43.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze. 1982. "An Experimental Analysis of Ultimatum Bargaining," 3 *Journal of Economic Behavior and Organization* 367–88.

Hart, Oliver, and John Moore. 2008. "Contracts as Reference Points," 123 *Quarterly Journal of Economics* 1–48.

Herweg, Fabian, and Klaus M. Schmidt. 2015. "Loss Aversion and Inefficient Renegotiation," 82 *Review of Economic Studies* 297–332.

Hoffman, David A., and Tess Wilkinson-Ryan. 2013. "The Psychology of Contract Precautions," 80 *University of Chicago Law Review* 395–445.

Holt, Charles A., and Susan K. Laury. 2002. "Risk Aversion and Incentive Effects," 92 *American Economic Review* 1644–55.

Hoppe, Eva I., and Patrick W. Schmitz. 2011. "Can Contracts Solve the Hold-Up Problem? Experimental Evidence," 73 *Games and Economic Behavior* 186–99.

Humphrey, Steven J., and Stefan Mondorf. 2021. "Testing the Causes of Betrayal Aversion," 198 *Economics Letters* 109663.

Kahneman, Daniel. 1992. "Reference Points, Anchors, Norms, and Mixed Feelings," 51 *Organizational Behavior and Human Decision Processes* 296–312.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1991. "Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias," 5 *Journal of Economic Perspectives* 193–206.

Klein, Benjamin. 1988. "Vertical Integration as Organizational Ownership: The Fisher Body-General Motors Relationship Revisited," 4 *Journal of Law, Economics, & Organization* 199–213.

Klein, Benjamin. 2000. "Fisher—General Motors and the Nature of the Firm," 43 *Journal of Law and Economics* 105–42.

Klein, Benjamin, Robert G. Crawford, and Armen A. Alchian. 1978. "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," 21 *Journal of Law and Economics* 297–326.

Koehler, Jonathan J., and Andrew D. Gershoff. 2003. "Betrayal Aversion: When Agents of Protection Become Agents of Harm," 90 *Organizational Behavior and Human Decision Processes* 244–61.

Koszegi, Botond. 2014. "Behavioral Contract Theory," 52 *Journal of Economic Literature* 1075–118.

Köszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences," 121 *Quarterly Journal of Economics* 1133–65.

Langer, Ellen J. 1975. "The Illusion of Control," 32 *Journal of Personality and Social Psychology* 311–28.

Larrick, Richard P., Katherine A. Burson, Jack B. Soll. 2007. "Social Comparison and Confidence: When Thinking You're Better than Average Predicts Overconfidence (and When it Does not)," 102 *Organizational Behavior and Human Decision Processes* 76–94.

Lewinsohn-Zamir, Daphna. 2012. "The Questionable Efficiency of the Efficient Breach Doctrine," 168 *Journal of Institutional and Theoretical Economics* 5–26.

Liebowitz, Stan J., and Stephen E. Margolis. 1995. "Path Dependence, Lock-in, and History," 11 *Journal of Law, Economics, & Organization* 205–26.

McAfee, R. Preston, John McMillan, and Michael D. Whinston. 1989. "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values," 104 *Quarterly Journal of Economics* 371–83.

Murphy, Ryan O., and Kurt A. Ackermann. 2014. "Social Value Orientation: Theoretical and Measurement Issues in the Study of Social Preferences," 18 *Personality and Social Psychology Review* 13–41.

Nalebuff, Barry. 2004. "Bundling as an Entry Barrier," 119 *Quarterly Journal of Economics* 159–87.

Perugini, Marco, Marcello Gallucci, Fabio Presaghi, and Anna Paola Ercolani. 2003. "The Personal Norm of Reciprocity," 17 *European Journal of Personality* 251–83.

Polinsky, A. Mitchell. 1983. "Risk Sharing Through Breach of Contract Remedies," 12 *Journal of Legal Studies* 427–44.

Posner, Richard A. 2009. "Let Us Never Blame a Contract Breaker," 107 *Michigan Law Review* 1349–63.

Presson, Paul K., and Victor A. Benassi. 1996. "Illusion of Control; A Meta-Analytic Review," 11 *Journal of Social Behavior and Personality* 493–510.

Rabin, Matthew. 1993. "Incorporating Fairness into Game Theory and Economics," 83 *American Economic Review* 1281–302.

Savage, Leonard J. 1954. *The Foundations of Statistics*. New York, NY: Wiley.

Schmitt, Manfred, Anna Baumert, Mario Gollwitzer, and Jürgen. Maes. 2010. "The Justice Sensitivity Inventory: Factorial Validity, Location in the Personality Facet Space, Demographic Pattern, and Normative Data," 23 *Social Justice Research* 211–38.

Schwartz, Barry, Andrew Ward, John Monterosso, Sonja Lyubomirsky, Katherine White, and Darrin R. Lehman. 2002. "Maximizing Versus Satisficing: Happiness is a Matter of Choice," 83 *Journal of Personality and Social Psychology* 1178–97.

Selten, Reinhard. 1967. "Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments," in Ernst Sauermann, ed., *Beiträge zur experimentellen Wirtschaftsforschung*. Tübingen: Mohr.

Shaw, Alex, Anam Barakzai, and Boaz Keysar. 2019. "When and Why People Evaluate Negative Reciprocity as More Fair Than Positive Reciprocity," 43 *Cognitive Science* e12773.

Sloof, Randolph, Joep Sonnemans, and Hessel Oosterbeek. 2004. "Specific Investments, Holdup, and the Outside Option Principle," 48 *European Economic Review* 1399–410.

Tversky, Amos, and Daniel Kahneman. 1991. "Loss Aversion in Riskless Choice: A Reference-Dependent Model," 106 *Quarterly Journal of Economics* 1039–61.

Van de Ven, Niels, and Marcel Zeelenberg. 2011. "Regret Aversion and the Reluctance to Exchange Lottery Tickets," 32 *Journal of Economic Psychology* 194–200.

Wilkinson-Ryan, Tess, and Jonathan Baron. 2009. "Moral Judgment and Moral Heuristics in Breach of Contract," 6 *Journal of Empirical Legal Studies* 405–23.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York, London: Free Press; Collier Macmillan.

Zamir, Eyal, and Doron Teichman. 2018. *Behavioral Law and Economics*. Oxford: Oxford University Press.

Zeelenberg, Marcel, and Jane Beattie. 1997. "Consequences of Regret Aversion 2: Additional Evidence for Effects of Feedback on Decision Making," 72 *Organizational Behavior and Human Decision Processes* 63–78.

Zeelenberg, Marcel, Jane Beattie, Joop Van der Pligt, and Nanne K. De Vries. 1996. "Consequences of Regret Aversion: Effects of Expected Feedback on Risky Decision Making," 65 *Organizational Behavior and Human Decision Processes* 148–58.