Five Essays on Experimental Economics

Inauguraldissertation

zur Erlangung des Doktorgrades
durch die
Wirtschaftswissenschaftliche Fakultät
der Georg-August-Universität Göttingen

vorgelegt von
Asri Özgümüs
aus Göttingen

Göttingen, 2021

Erstgutachterin: Prof. Dr. Claudia Keser
Zweitgutachter: Prof. Dr. Kilian Bizer

Weiteres Mitglied der Prüfungskommission: Jun.-Prof. Dr. Holger A. Rau

Contents

1 Introduction

This dissertation consists of five essays contributing to a better understanding of three fundamental lines of research in behavioral economics: trust and cooperation in social interactions (essays in Appendix A.1–A.3), gender discrimination (essay in Appendix A.4), and decision making under risk and uncertainty (essay in Appendix A.5). All essays are novel contributions and based on the experimental economics method. In some of the essays, game-theoretical solutions are provided and experimental findings are related to behavioral theories.

Trust and cooperation in social interactions
In section 2.1, I summarize an essay on the behavior of trustors and trustees in the framework of an investment game (Berg, Dickhaut, and McCabe, 1995) where the trustee's action space is extended by the opportunity to take money from the trustor, in addition to the amount received. This makes trustors more vulnerable toward trustees in that trustees can do additional harm to trustors beyond not returning any money. This important aspect has not been featured in the investment game to date, but is of great relevance in practice. Imagine, for example, a firm sharing its confidential employee data with a services company; those data could be sold to a third party. Will trustors place trust in a framework where doing so provides the trustee with the opportunity to misuse, in addition to not reciprocate, trust? If so, will trustees make use of this opportunity?

In section 2.2, I summarize an essay on the cooperative behavior of issuers and credit rating agencies (CRAs) in a simplified rating-market setup where regulatory reforms are strictly implemented. The reforms concern the prevalent *issuer-pays* compensation model of CRAs, which allows issuers and CRAs to collude on an outcome where CRAs offer overly optimistic ratings and issuers choose the most favorable among the solicited ratings (e.g., White, 2010; Jiang, Stanford, and Xie, 2012; Skreta and Veldkamp, 2009; Bolton, Freixas, and Shapiro, 2012). Numerous reforms have been suggested in both theory and practice. We develop a simple game-theoretical model that captures the main aspects of the repeated interaction between an issuer and a

CRA under those reforms that theoretically seem promising. We then study how issuers and CRAs behave in a controlled laboratory environment, allowing a maximum of control of extraneous variables that might promote cooperative behavior in both theory and practice. Will issuers and CRAs be able to establish cooperation even in such a strictly regulated compensation context?

In section 2.3, I summarize an essay on the strategic interaction of consumers and sellers in a Pay-What-You-Want (PWYW) framework (Schmidt, Spann, and Zeithammer, 2015). In such a framework of participative pricing (Kim, Natter, and Spann, 2009), consumers are allowed to freely choose the price they are willing to pay for a seller's product, including zero. Cooperation between consumers and PWYW sellers arises when consumers pay prices that are sufficiently high and sellers decide to offer their products under the PWYW pricing mechanism rather than conventional price setting (e.g., Schmidt, Spann, and Zeithammer, 2015). In many practical applications of PWYW pricing, consumers choose which price to pay not in isolation, but in the presence of other consumers. As an illustration, imagine a theater visitor who is standing at the counter and surrounded by other visitors while making the payment decision. In such an environment, cooperation between consumers and PWYW sellers might be more pronounced as voluntary payments are largely based on consumers' fairness concerns (e.g., Kim, Natter, and Spann, 2009) and fairness concerns become stronger while being observed, which can be rationalized by image concerns (e.g., Bénabou and Tirole, 2006; Ellingsen and Johannesson, 2008; Andreoni and Bernheim, 2009). In our experiment, we exogenously vary the environment in which consumers make voluntary payments to PWYW sellers in order to find out in which environment voluntary payments are maximized. In particular, are consumers' payments to a PWYW seller higher depending on whom they are surrounded with (e.g., Schlüter and Vollan, 2015), by whether or not their payments are observable (e.g., Gneezy et al., 2012), or a combination of both?

Gender discrimination

In section 2.4, I summarize an essay on gender discrimination in the context of university teaching evaluations. In this context, it has been shown empirically that male raters discriminate against female instructors in that they assign systematically higher performance ratings to male as compared to female instructors (e.g., Boring, 2017; Mengel, Sauermann, and Zölitz, 2019). The evaluation differences observed in these studies can neither be explained by a selection bias (e.g., Heckman, 2010)

of raters nor by differences in the actual teaching performance of instructors. Yet, it is difficult to identify a bias as the cause of the evaluation differences. Other factors might explain the evaluation differences, which are difficult to exclude when studied outside a controlled environment. We thus employ a laboratory experiment that enables us to hold constant any factors that potentially enter and distort the evaluation process of raters. This allows us to identify a bias as the cause of the observed differences in evaluations, if any. Do performance ratings toward female and male instructors become gender-unbiased in the most reduced context without any personal interaction with instructors and potential other factors distorting the evaluations of raters?

Decision making under risk and uncertainty

In section 2.5, I summarize an essay on individual as well as social decision making under temporally delayed risks. In experimental economics, there is selective evidence on temporal delay effects in the domains of individual decision making under risk (e.g., Abdellaoui, Diecidue, and Öncüler, 2011; Pan, Web, and Zank, 2019) as well as social decision making under risk, i.e., trust (e.g., Neo et al., 2013; Ederer and Schneider, forthcoming). In both domains the studies differ in various methodological aspects and so do not allow for a unifying analysis of decision making under temporally delayed risks, respectively. In addition, the preceding evidence of temporal delay effects on individual decisions under risk is incomplete in that previous evidence has focused on individual known risk, i.e., situations where the probability distribution underlying a risky lottery is known. However, most scenarios in reality are characterized by individual unknown (ambiguous) risk, i.e., situations where the probability distribution underlying a risky lottery is unknown to the decision maker (Ellsberg, 1961; Trautmann and van de Kuilen, 2015). In order to get a comprehensive understanding of the impact of temporal distance in decisions under individual risk, both known and unknown, as well decisions under social risk, we use the psychological framework of construal level theory (CLT; Liberman and Trope, 1998; Trope and Liberman, 2000, 2003, 2010). To what extent is CLT able to unify and describe the observed behavioral patterns of individual and social decision making under temporally delayed risks?

2 Summary of Essays

2.1 Trust, Vulnerability and Trustworthiness

(coauthored with Claudia Keser)

published in
Economics Letters

Motivation: In the experimental paradigm of Berg, Dickhaut, and McCabe (1995), we extent the trustee's action space by the opportunity to take money from the trustor, in addition to the amount received. This makes trustors more vulnerable toward trustees in that trustees can do additional harm to trustors beyond not returning any money. Using the concept of backward induction leads to the same subgame perfect equilibrium solution (Selten, 1965) in the setting with the extended action space and the original investment game by Berg, Dickhaut, and McCabe (1995). Nevertheless, it is likely that the behavior of trustors and trustees differs in the two settings. In particular, behavior is based on expectations and different action spaces invoke different expectations, as is demonstrated in related contexts of dictator games (e.g., List, 2007; Bardsley, 2008). In these contexts, dictator players give significantly less amounts of money to recipients when their action space of giving or keeping money is extended by the option to take money from the recipients' endowments. A main reason for this finding is that dictator players form their expectations about what is appropriate to give based on their available actions (List, 2007; Bardsley, 2008). For example, choosing to give nothing in a setting, where the alternatives are to give or to take money from the recipient, might be considered as being fair, since no money is taken from the recipient. However, choosing to give nothing, when the only other choice is to give money to the recipient, might not be considered as being fair, since it is the worst option for the recipient.

We put this line of argumentation into a context where social interaction is involved. Specifically, we study if the same line of argumentation applies in the context of an investment game, where the trustee's decision is technically similar to the decision in a dictator game, but preceded by the decision of a trustor. This might have two important implications for the behavior of trustees, allowing us to shed light on the role of expectations in a more realistic environment. First, trustors who expect trustees to weigh the fairness of their actions based on their available actions might invest less in the game with the larger action space, because in keeping with the line of argumentation above, they shall expect lower amounts returned from trustees. The lower investments by trustors might then indeed lead to lower amounts returned by trustees. Second, trustors might expect trustees who have been given a positive amount to attribute intentional kindness to their decisions. Intentional kindness is a key determinant of positively reciprocal behavior and can be inferred from which choice has been made taking into account the available alternatives (e.g., Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk, Fehr, and Fischbacher, 2008). In our context, investing positive amounts despite being highly vulnerable toward trustees might be interpreted as being intentionally kind. Hence, trustors who invest positive amounts and expect trustees to attribute intentional kindness to their actions might then decide to not invest less in the game with the larger action space, because they shall expect positively reciprocal behavior from trustees. This in turn might increase the amounts returned by trustees. In a controlled environment, we try to explain trustors' and trustees' behavior in the investment with the larger action space by drawing on the two considerations just made.

Methods: We design and implement a between-subjects design with subjects being randomly assigned to one of two experimental conditions: BASE and TAKE. In BASE, subjects take part in an investment game that is similar to the experimental paradigm by Berg, Dickhaut, and McCabe (1995). Both trustors and trustees receive an initial endowment of 12 Euros, half of which may be used by trustors to invest. Any amount invested by trustors is quadrupled so that trustees receive the quadrupled amount invested by trustors. Trustees may keep the amount received, or return part or all of it to the trustors. Payoffs are such that the trustor receives the initial endowment minus the amount invested plus the amount returned, and the trustee receives the initial endowment plus the quadrupled amount invested minus the amount returned. TAKE is identical to BASE in every respect, except for one feature: in addition to the quadrupled amount received, trustees may take some positive amount from the trustors' remaining endowments, at most, however, up to

the amount invested. This alters the trustee's payoff function as follows: instead of subtracting the amount returned from the trustee's payoff, positive amounts taken, if any, might be added. In both experimental conditions, we ask trustors to state the amount of money they expect to be returned or taken away by trustees as well as the amounts they would consider as fair.

Main findings: Our main findings indicate that trustors invest significantly lower amounts in TAKE as compared to BASE, which can neither be explained by trustors' expectations nor by their fairness considerations in our experiment. The amounts returned by trustees in terms of a relative-payback measure remain largely unaffected by the trustors' decisions, although trustees, when given the opportunity, refrain from taking away even a single incidence from trustors. We may interpret this result as intention-based reciprocity, as investing positive amounts despite a very high risk of being exploited can serve the trustee as a strong signal of the trustor's intention to enter a reciprocal-trust relationship (McCabe, Rigdon, and Smith, 2003).

2.2 An Experimental Investigation of Rating-Market Regulation

(coauthored with Claudia Keser, Emmanuel Peterlé, and Martin Schmidt)

published in

Journal of Economic Behavior & Organization

Motivation: We study in a controlled environment the effectiveness of a regulatory reform intended to prevent cooperation between issuers and credit rating agencies (CRAs) under the prevalent *issuer-pays* compensation model of CRAs. The issuer-pays model allows issuers of financial instruments to observe the credit ratings, which CRAs assign to their products, before deciding whether or not to purchase the ratings and thus make them public to potential investors (e.g., Securities and Exchange Commission, 2008). CRAs receive over 90 percent of their income from issuer fees (e.g., Darcy, 2009), which makes them highly dependent on those fees and hence, puts into question their autonomy in the provision of accurate ratings. Evidence from financial economics shows that the issuer-pays model makes CRAs to inflate ratings, i.e., offer overly optimistic ratings in order to attract business (e.g., White, 2010; Jiang, Stanford, and Xie, 2012), and issuers to shop for ratings, i.e., solicit multiple ratings from different CRAs and choose the most favorable one (e.g., Skreta and Veldkamp, 2009; Bolton, Freixas, and Shapiro, 2012).

In order to increase the overall quality, i.e., accuracy of ratings, numerous regulatory reforms have been suggested. In 2008, Andrew Cuomo, then Attorney General and today's governor of New York, reached a three-year agreement with today's most famous CRAs, which required issuers of structured finance products to pay the issuer fee up front, i.e., before a CRA conducts its initial analysis (New York State Office of the Attorney General, 2008). This reform was intended to make CRAs less dependent on the question of receiving the issuer fee or not and thus to increase their autonomy in providing accurate credit ratings. Even though the success of this

reform was empirically not evaluated, it is doomed to fail from a theoretical point of view. Issuers might consider up-front payments as sunk costs and thus be willing to pay, ex post, for an agreement with the CRA not to publish a bad rating (Bolton, Freixas, and Shapiro, 2012). A much better solution could be a reform that prohibits rating shopping by enforcing CRAs to automatically disclose any rating requested and paid for by an issuer (Bolton, Freixas, and Shapiro, 2012).
While this seems promising in theory, we might expect the real behavior of issuers and CRAs to differ, in particular when interaction is repeated and rating markets are highly concentrated: issuers and CRAs are likely to have an interest in a long-term cooperation where issuers are interested in good ratings, which CRAs might be willing to supply since they are interested in future rating requests (e.g., Ashcraft, Goldsmith-Pinkham, and Vickery, 2010; Griffin and Tang, 2021; Frenkel, 2015). If this cooperative interest is strong enough, the suggested reforms are unlikely to prevent CRAs from inflating ratings, even when rating shopping is impossible and rating disclosure is enforced. We shed light on this controversy by conducting an experiment in a controlled laboratory environment.

Methods: We introduce a simple game-theoretical model that captures the main aspects of the repeated interaction between an issuer and a monopolistic CRA under strict regulatory reforms. The reforms concern the issuer-pays compensation model of CRAs and include no communication and no possibilities for hidden agreements, up-front payments of rating-fees (as suggested and implemented by Andrew Cuomo), direct publication of any rating that has been requested (as suggested by Bolton, Freixas, and Shapiro, 2012), and in addition, financial penalties for inaccurate ratings. In a between-subjects design, we vary the level of the financial penalty (low vs. high) that the CRA incurs by providing inaccurate ratings. In each repetition of the finite game, an issuer may request a rating of its financial product from a monopolistic CRA. If a rating is being requested, the CRA privately observes the true quality of the issuer's product and decides whether to provide and publish either an accurate or an inaccurate rating. CRAs obtain the full issuer fee if they provide accurate ratings, and are penalized for providing inaccurate ratings. For each rating that is inaccurate, a CRA is deducted 10 percent of its income in the low penalty condition and 50 percent of its income in the high penalty condition. The game is parametrized such that the subgame perfect equilibrium is for the CRA to always report accurately and knowing this, for the issuer to never request a rating.

Main findings: We observe that, against the subgame-perfect solution of the rating-market game, issuers request ratings in about 75% of the cases. CRAs cooperate and provide overly optimistic ratings in more than 50% of the cases, independent of the level of the financial penalty for inaccurate ratings. Moreover, regression results indicate that providing an overly optimistic rating significantly increases the likelihood of another cooperative outcome with the issuer, while providing an overly pessimistic rating significantly decreases the likelihood of another cooperative outcome with the issuer. Our findings can be explained by the chain store paradox of Selten (1978), where both players pursuing a "deterrence strategy" (in our case not requesting a rating and not providing an overly optimistic rating) in some periods of the finitely repeated game might help to establish a mutually beneficial outcome in the long run. Cooperation between issuers and CRAs is therefore to be expected also in reality, where the cooperative interests of issuers and CRAs might be promoted by many other factors.

2.3 Who is Watching me? Disentangling Audience and Interpersonal Closeness Effects in a Pay-What-You-Want Context

(coauthored with Elisa Hofmann, Michael E. Fiagbenu, Amir M. Tahamtan, and Tobias Regner)

published in

Journal of Behavioral and Experimental Economics

Motivation: We study if consumers in a Pay-What-You-Want (PWYW) context make higher voluntary payments when their payments are observable by close others. In a PWYW context, buyers can freely choose the price to pay for a seller's product. Hence, the seller's income is largely determined by the fairness preferences of its buyers (e.g., Kim, Natter, and Spann, 2009). It is well acknowledged in economic research that being observed by an audience increases buyers' concerns for fairness, which can be rationalized by image concerns (e.g., Bénabou and Tirole, 2006; Ellingsen and Johannesson, 2008; Andreoni and Bernheim, 2009). Related studies suggest that the more relevant (socially close) the observing audience is, the stronger the image concerns and consequently, the higher the voluntary payments being made. Most of these studies, however, vary at the same time the observability of payments and the social closeness of the observing audience, which leaves the question open which of the two factors is actually responsible for the increase in buyers' payments. For example, it might be sufficient to be surrounded without being observed by a close audience to increase payments. We, therefore, design and implement an experiment tailored to disentangle the causal effects of the observability of payments and the social closeness of the observing audience on voluntary payments.

Methods: We use the experimental paradigm by Schmidt, Spann, and Zeithammer (2015) and let buyers in groups of three repeatedly make voluntary payments to a PWYW seller. Each buyer makes individual payment decisions knowing that she or he is connected with two other buyers, who define the surrounding audience. In our 2x2 between-subjects design, we vary the observability of payments and the interpersonal closeness of the surrounding audience. Payments are observable means that a buyer's payment decision is publicly known to the other group members. Interpersonal closeness is induced by using a method from social psychology, where buyers in a group are asked to stand up and look into each other's eyes before the start of the experiment.

Main findings: We compare buyers' average payments in the four treatments and find that being surrounded by close others as well as payments being observable by close others increases voluntary payments. While both effects are positive and additive in total, they do not reinforce each other. That is, we do not find any interaction effect. This finding leaves open an important question on the importance of image concerns in close audiences. In particular, image concerns might become stronger with the social closeness of the observing audience only up to a certain threshold. When the observing audience becomes socially too close such as in the case of friends or family, image concerns might become less important to the decision maker since her or his generosity might be well-acknowledged in these audiences already.

2.4 Gender Bias in the Evaluation of Teaching Materials

(coauthored with Holger A. Rau, Stefan T. Trautmann, and Christian König-Kersting)

published in

Frontiers in Psychology

Motivation: We provide clean and robust evidence for gender biases in the evaluation of university teaching materials. Recent empirical evidence in economics suggest gender biases in evaluation processes in academia, often driven by male evaluators in favor of male instructors (e.g., Boring, 2017; Mengel, Sauermann, and Zölitz, 2019). Even though these studies do not investigate gender differences in evaluations in a controlled environment, they can exclude justifying explanations such as differences in the teaching performance of female and male instructors. In particular, they achieve this by demonstrating that the observed differences in evaluations do not correlate with any teaching effectiveness measures. However, the gender differences observed in these studies might not necessarily point at a gender bias, which is much more difficult to identify. Other explanatory factors might play a role, which are difficult to exclude when not studied in a controlled environment. For example, subjective in-class experiences of evaluators or subjective experiences with instructors (e.g., MacNell, Driscoll, and Hunt, 2015) might play an important role. Alternatively, there might be a fundamental gender bias unrelated to any factual class or instructor experience. Our aim is to separate these two explanations and study in a controlled environment if performance ratings toward female and male instructors become gender-unbiased, if any subjective experiences of female and male evaluators are being eliminated.

Methods: In a laboratory setting, we present subjects with two sets of lecture slides from an economics course in academia. The two sets of slides are identical in content

but differ in their layout. Between subjects, we vary the gender of the instructor by randomly assigning either a female or a male first name to the title page of the two slide sets. The surname is kept constant as the common German name "Müller." Subjects browse through the two sets of slides and evaluate them along several dimensions. Six items concern the general quality of the slides, and three further items concern the comparison of the two layouts presented. The latter only serves to provide the study background for subjects. We sum up the six quality-related items to an index of the evaluators' perceived quality of the teaching material, which we then compare between the induced gender. In order to generalize our findings from this study to a more heterogeneous subject pool (mainly non-economists, partly non-academic background, and partly current students), we conduct another study with shorter and less technical slides online on Amazon Mechanical Turk (MTurk), keeping every other detail constant between the two studies.

Main findings: In two studies, one with a German student sample and the other with a more heterogeneous subject pool, we find substantial lower evaluations of female-named instructors as measured by our quality index. In accordance with previous studies, the lower evaluations are driven by male evaluators with past experiences made in university classes. No gender biases are observed for evaluators without a college degree, i.e., without past experiences made in university classes. While in the student sample female evaluators seem rather gender-unbiased in their evaluations, female evaluators in the more heterogeneous subject pool tend to evaluate female-named instructors more favorably than male-named instructors, if they own a college degree. Overall, our findings imply that subtle biases toward gender exist even in the most reduced contexts without any personal interaction and potential other factors distorting the raters' evaluations. However, the gender bias we observe is not uniform in that study experiences, in general, seem to play an important role for the bias to emerge.

2.5 Delayed Risks in Individual and Social Decisions

(coauthored with Holger A. Rau and Stefan T. Trautmann)

working paper

Motivation: We aim to get a comprehensive understanding of individual and social behavior under temporally delayed risks using the psychological framework of construal level theory (CLT; Liberman and Trope, 1998; Trope and Liberman, 2000, 2003, 2010). While preceding evidence on temporal delay effects in economics is rather selective and differs in many methodological aspects, CLT allows us to perform a unifying analysis of individual and social decisions under risk in both the near and the distant future. CLT proposes that individuals trade off an outcome's *desirability* with its *feasibility*, and weight the outcome *desirability* relatively lower than the outcome *feasibility* in the near future and vice versa in the more distant future. The underlying reason is that outcome *desirability* and outcome *feasibility* are mentally represented on different construal levels (Liberman and Trope, 1998; Sagristano, Trope, and Liberman, 2002; Trope and Liberman, 2003). In the context of gambling decisions, it has been shown that the outcome *desirability* is associated with the payoff of a risky gamble and the outcome *feasibility* is associated with the probability of winning of a risky gamble (Sagristano, Trope, and Liberman, 2002). CLT thus implies, in an economic sense, that individuals weigh the payoff of a risky lottery relatively lower than its probability of winning in the near future and relatively higher than its probability of winning in the more distant future.

In our experiment, we test the implications of CLT along three dimensions of risk: individual known risk, individual unknown (ambiguous) risk, and social risk. In the dimensions of individual known and individual ambiguous risks, subjects value lotteries characterized by either a low probability of winning a high payoff (i.e., "low–*feasibility*–high–*desirability* lottery") or a high probability of winning a low

payoff (i.e., "high–*feasibility*–low–*desirability* lottery"). Whereas in the dimension of individual known risk the underlying probability distribution of winning either a high or a low payoff is known, it is unknown in the dimension of individual ambiguous risk (Ellsberg, 1961; Trautmann and van de Kuilen, 2015). In both domains, we study the implications of CLT by testing if individuals value low–*feasibility*–high–*desirability* lotteries higher in the more distant future as compared to the near future, and if individuals value high–*feasibility*–low–*desirability* lotteries higher in the near future as compared to the more distant future. Our setting also allows us to study the implications of CLT for the ambiguity attitudes of individuals (Fox and Tversky, 1995; Trautmann and van de Kuilen, 2015). For this purpose, we hold constant the level of *feasibility* and *desirability*, respectively, and compare subjects' valuations of risky lotteries with their ambiguous counterparts. That is, we compare subjects' valuations of risky low–*feasibility*–high–*desirability* lotteries with their valuations of ambiguous low–*feasibility*–high–*desirability* lotteries, and subjects' valuations of risky high–*feasibility*–low–*desirability* lotteries with their valuations of ambiguous high–*feasibility*–low–*desirability* lotteries.

We complement our analysis of individual delayed risks with the analysis of socially delayed risks. That is, we analyze trustors' decisions in an investment game (Berg, Dickhaut, and McCabe, 1995) when they are confronted with a delayed decision of the trustee. In particular, we let trustors make investment decisions knowing that trustees will make their decision either in the near or the distant future.

Methods: In a between-subjects design, we identify the causal effects of temporal delay by varying the point in time when the risks involved in the individual and the social context are resolved. The risks are resolved either immediately (*Early* condition) or two weeks later (*Late* condition). Except for the timing of the resolution of risks, both experimental conditionals are exactly identical in every respect. Each of the two conditions consists of two experimental sessions.

In the first experimental session, we first implement a lottery task in order to elicit subjects' valuation of individual risk. Specifically, we elicit in an incentive-compatible way subjects' willingness to pay (WTP) for two risky and two ambiguous lotteries presented to subjects as a pair, each. Each pair of lotteries consists of a lottery that is characterized by a high *feasibility* and a low *desirability*, and another one that is characterized by a low *feasibility* and a high *desirability*. To ensure the comparability of subjects' WTPs between lotteries, all lotteries are parametrized such that they have the same expected value. In the second part of the first experimental session, we implement an investment game (Berg, Dickhaut, and McCabe, 1995) in order to

elicit subjects' valuations of social risk. Subjects make decisions in the role of first movers, knowing that the decision of the second mover will be made either immediately afterwards (*Early* condition) or two weeks later (*Late* condition).
In order to eliminate confounds such as time-preference effects for monetary outcomes, we pay out subjects always at the end of the second experimental session. Regarding individual risk, we make within-subject comparisons of high–*feasibility*–low-*desirability* and low–*feasibility*–high–*desirability* lotteries. We make this comparison in the domains of both risk and ambiguity in order to test the CLT implications in the near future, respectively. We also make within-subject comparisons of risky high–*feasibility*–low–*desirability* lotteries and ambiguous high–*feasibility*–low-*desirability* lotteries, and within-subject comparisons of risky low–*feasibility*–high-*desirability* lotteries and ambiguous low–*feasibility*–high-*desirability* lotteries in order to test the CLT implications for individual ambiguity attitudes in the near future. Furthermore, we make between-subject comparisons of WTP differences in high–*feasibility*–low–*desirability* and low–*feasibility*–high–*desirability* lotteries in order to identify the causal effect of temporal delay on individual risk. Regarding social risk, we make between-subject comparisons of first-mover decisions in order to identify the causal effect of temporal delay on social risk.

Main findings: In the case of individual risk, we find that subjects value high–*feasibility*–low–desirability lotteries always higher than low–*feasibility*–high–*desirability* lotteries. Under ambiguity, this gap decreases when the resolution of the lottery occurs in the distant future. We find no temporal distance effect for the valuation of lotteries with known risk. Moreover, we find that subjects are always ambiguity averse, which becomes less pronounced when low–*feasibility*–high–*desirability* lotteries are resolved in the distant future. In the case of social risk, we find that time distance lowers trusting behavior as trustors correctly anticipate that reciprocity is lower when trustees decide in the future. In sum, we find some support for temporal distance effects on decisions under risk, which are consistent with the psychological predictions of CLT. However, these effects turn out to be relatively small, rendering no difference in the qualitative interpretation of findings between non-delayed and delayed contexts of the resolution of risks. Consequently, we conclude that the external validity of qualitative results of lab experimental studies on decision under risk will not be severely affected if temporal resolution in the lab differs from temporal resolution outside the lab.

3 Conclusion

The five essays included in this dissertation have covered different areas of behavioral economics using the experimental method. They let us draw several lessons, three of which will be addressed in the following.

First, from data gathered in the experimental laboratory, important implications may be derived for the behavior of economic agents in reality, even though lab-experimental data is associated with weak external validity (e.g., Levitt and List, 2007). For example, regulatory reforms that theoretically prove promising were not able to prevent cooperation of economic agents in a simplified but strict laboratory environment (section 2.2). This gives policy makers an indication to rethink implementing their intended reforms in practice, where institutional environments are less strict and the interaction of economic agents is much more natural and rich.

Second, tailored experimental designs enable the identification of partial effects (section 2.3) and subtle biases (section 2,4) that would otherwise be difficult to achieve. In this sense, laboratory data can complement the external validity of secondary data that is being used for empirical analyses.

Third and lastly, considering psychological frameworks in the analysis of economic decisions can help to generalize research findings within economics over different contexts of decision making (section 2.5).

Bibliography

Abdellaoui, M., Diecidue, E., and Öncüler, A., 2011. Risk Preferences at Different Time Periods: An Experimental Investigation. *Management Science*, 57(5), 975-987.

Andreoni, J. B. and Bernheim, D., 2009. Social Image and the 50-50 Norm: A Theoretical and Experimental Analysis of Audience Effects. *Econometrica*, 77(5), 1607–1636.

Ashcraft, A. B., Goldsmith-Pinkham, P., and Vickery, J. I., 2010. MBS Ratings and the Mortgage Credit Boom. Federal Reserve Bank of New York Staff Report 449.

Bardsley, N., 2008. Dictator Game Giving: Altruism or Artefact? *Experimental Economics*, 11(2), 122-133.

Bénabou, R. and Tirole, J., 2006. Incentives and Prosocial Behavior. *American Economic Review*, 96(5), 1652–1678.

Berg, J., Dickhaut, J., and McCabe, K., 1995. Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 10(1), 122–142.

Bolton, P., Freixas, X., and Shapiro, J., 2012. The Credit Ratings Game. *Journal of Finance*, 67(1), 85-111.

Boring, A., 2017. Gender Biases in Student Evaluations of Teaching. *Journal of Public Economics*, 145, 27–41.

Darcy, D., 2009. Credit Rating Agencies and the Credit Crisis: How the 'Issuer Pays' Conflict Contributed and what Regulators might do about it. *Columbia Business Law Review*, 2009(2), 605–668.

Dufwenberg, M. and Kirchsteiger, G., 2004. A Theory of Sequential Reciprocity. *Games and Economic Behavior*, 47(2), 268–298.

Ederer, F. and Schneider, F. (forthcoming). Trust and Promises over Time. *American Economic Journal: Microeconomics*.

Ellingsen, T. and Johannesson, M., 2008. Pride and Prejudice: The Human Side of Incentive Theory. *American Economic Review*, 98(3), 990–1008.

Ellsberg, D., 1961. Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics*, 75(4), 643-669.

Falk, A., Fehr, E., and Fischbacher, U., 2008. Testing Theories of Fairness – Intentions Matter. *Games and Economic Behavior*, 62(1), 287-303.

Fox, C. R. and Tversky, A., 1995. Ambiguity Aversion and Comparative Ignorance. *Quarterly Journal of Economics*, 110(3), 585-603.

Frenkel, S., 2015. Repeated Interaction and Rating Inflation: A Model of Double Reputation. *American Economic Journal: Microecononomics*, 7(1), 250–280.

Gneezy, A., Gneezy, U., Riener, G., and Nelson, L. D., 2012. Pay-What-You-Want, Identity, and Self-Signaling in Markets. *Proceedings of the National Academy of Sciences of the United States of America*, 109(19), 7236–7240.

Griffin, J.M. and Tang, D.Y., 2012. Did Subjectivity Play a Role in CDO Credit Ratings? *Journal of Finance*, 67(4), 1293–1328.

Heckman, J. J., 2010. Selection Bias and Self-Selection. In: S. N. Durlauf and L. E. Blume (eds), Microeconometrics, 242-266.

Jiang, J.X., Stanford, M.H., and Xie, Y., 2012. Does it Matter who Pays for Bond Ratings? Historical Evidence. *Journal of Financial Economics*, 105(3), 607–621.

Kim, J.-Y., Natter, M., and Spann, M., 2009. Pay What You Want: A New Participative Pricing Mechanism. *Journal of Marketing*, 73(1), 44–58.

Levitt, S. D. and List, J. A., 2007. What do Laboratory Experiments Measuring Social Preferences Reveal about the Real World? *Journal of Economic Perspectives*, 21(2), 153-174.

Liberman, N. and Trope, Y., 1998. The Role of Feasibility and Desirability Considerations in Near and Distant Future Decisions: A Test of Temporal Construal Theory. *Journal of Personality and Social Psychology*, 75(1), 5-18.

List, J. A., 2007. On the Interpretation of Giving in Dictator Games. *Journal of Political Economy*, 115(3), 482–493.

MacNell, L., Driscoll, A., and Hunt, A. N., 2015. What's in a Name: Exposing Gender Bias in Student Ratings of Teaching. *Innovative Higher Education*, 40(4), 291–303.

McCabe, K. A., Rigdon, M. L., and Smith, V. L., 2003. Positive Reciprocity and Intentions in Trust Games. *Journal of Economic Behavior & Organization*, 52(2), 267–275.

Mengel, F., Sauermann, J., and Zölitz, U., 2019. Gender Bias in Teaching Evaluations. *Journal of the European Economic Association*, 17(2), 535–566.

New York State Office of the Attorney General, 2008. Attorney General Cuomo Announces Landmark Reform Agreements With The Nation's Three Principal Credit Rating Agencies. Available at: http://www.ag.ny.gov/press-release/attorney-general-cuomo-announces-landmark-reform-agreements-nations-three-principal (Retrieved on March 01, 2021).

Neo, W. S., Yu, M., Weber, R. A., and Gonzalez, C., 2013. The Effects of Time Delay in Reciprocity Games. *Journal of Economic Psychology*, 34, 20–35.

Pan, J., Webb, C. S., and Zank, H. 2019. Delayed Probabilistic Risk Attitude: A Parametric Approach. *Theory and Decision*, 87(2), 201-232.

Rabin, M., 1993. Incorporating Fairness into Game Theory and Economics. *American Economic Review*, 83(5), 1281-1302.

Sagristano, M. D., Trope, Y. and Liberman, N., 2002. Time-Dependent Gambling: Odds Now, Money Later. *Journal of Experimental Psychology: General*, 131(3), 364-376.

Schlüter, A. and Vollan, B., 2015. Flowers and an Honour Box: Evidence on Framing Effects. *Journal of Behavioral and Experimental Economics*, 57, 186–199.

Schmidt, K. M., Spann, M., and Zeithammer, R., 2015. Pay What You Want as a Marketing Strategy in Monopolistic and Competitive Markets. *Management Science*, 61(6), 1217–1236.

Securities and Exchange Commission, 2008. Summary Report of Issues Identified in the Commission Staff's Examinations of Select Credit Rating Agencies. Available at: https://www.sec.gov/news/studies/2008/craexamination070808.pdf (Retrieved on March 01, 2021).

Selten, R., 1965. Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. *Zeitschrift für die gesamte Staatswissenschaft*, 121(2), 301-324.

Selten, R., 1978. The Chain Store Paradox. *Theory and Decision*, 9 (2), 127–159.

Skreta, V. and Veldkamp, L., 2009. Ratings Shopping and Asset Complexity: A Theory of Ratings Inflation. *Journal of Monetary Economics*, 56(5), 678–695.

Trautmann, S.T. and van de Kuilen, G., 2015. Ambiguity Attitudes. In: G. Keren and G. Wu (eds), The Wiley Blackwell Handbook of Judgment and Decision Making 1, 89-116.

Trope, Y. and Liberman, N., 2000. Temporal Construal and Time-Dependent Changes in Preference. *Journal of Personality and Social Psychology*, 79(6), 876-889.

Trope, Y. and Liberman, N., 2003. Temporal Construal. *Psychological Review*, 110(3), 403-421.

Trope, Y. and Liberman, N., 2010. Construal-Level Theory of Psychological Distance. *Psychological Review*, 117(2), 440-463.

White, L. J., 2010. Markets: The Credit Rating Agencies. *Journal of Economic Perspectives*, 24(2), 211–226.

Appendix

A.1 Trust, Vulnerability and Trustworthiness

(coauthored with Claudia Keser)

Abstract

In the investment game by Berg et al. (1995), we extend the trustee's action space by the opportunity to take money from the trustor, in addition to the amount received. Experimental findings indicate that this significantly reduces the trustors' investment. While the trustees' average payback relative to the amount received remains unaffected, we observe an important decrease in the relative frequency of zero returns. Furthermore, we do not observe a single incidence of money being taken from the trustor. We explain this result by intention-based social preferences.

Keywords: Trust game, Experiment

JEL Classifications: C91, C72, D63

Published:

Economics Letters, Vol. 163, February 2018, pp. 149-151.

DOI: https://doi.org/10.1016/j.econlet.2017.12.012

A.2 An Experimental Investigation of Rating-Market Regulation

(coauthored with Claudia Keser, Emmanuel Peterlé, and Martin Schmidt)

Abstract

We introduce a simple game-theoretical model that captures the main aspects of the repeated interaction between an issuer and a credit-rating agency. It involves up-front payments of issuer-fees and direct publication of requested ratings. Due to pecuniary injuries for untruthful ratings, the credit-rating agency should always report truthfully in the subgame perfect equilibrium. Knowing this, the issuer should never request a rating. Conducting laboratory experiments, we find that behavior significantly deviates from the equilibrium prediction in favor of a cooperative solution: issuers frequently do request ratings, which is often reciprocated with untruthful good ratings.

Keywords: Game theory, Laboratory experiments, Rating agencies, Regulation

JEL Classifications: C70, C9, G0

Highlights:

- Repeated CRA-issuer interaction is examined in a simple experimental game.
- CRAs have an incentive to build a reputation for leniency.
- Issuers and CRAs frequently cooperate at the detriment of naïve clients.
- Regulation requesting issuers' up-front payment of rating fees is doomed to fail.

Published:
Journal of Economic Behavior & Organization, Vol. 144, December 2017, pp. 78-86.
DOI: https://doi.org/10.1016/j.jebo.2017.09.022

A.3 Who is Watching me? Disentangling Audience and Interpersonal Closeness Effects in a Pay-What-You-Want Context

(coauthored with Elisa Hofmann, Michael E. Fiagbenu, Amir M. Tahamtan, and Tobias Regner)

Abstract

We disentangle by means of a laboratory experiment two relevant drivers of voluntary payments in Pay-What-You-Want settings: the effects of interpersonal closeness and an audience. Our 2x2 between-subjects design varies the interpersonal closeness of buyers and the observability of their payments by other buyers. This allows us to enrich the research on both drivers and identify whether the presence of close others (closeness effect), payment observability (audience effect), or the combination of both affects voluntary payments. We find that both effects separately are sufficient to increase voluntary payments. Payments are, on average, higher if they are observed by an audience and if buyers are acquainted with each other. While the effect of audience and interpersonal closeness on payments is additive in total, we do not find an interaction effect, if payments are observed by close others.

Keywords: Social preferences, Experiments, Social image concerns, Pay-What-You-Want, Interpersonal closeness

JEL Classifications: C91, D03, L11

Highlights:

- We experimentally investigate payment determinants in a Pay-What-You-Want context.
- Our design varies the interpersonal closeness of buyers (strangers vs. acquaintances).

- We also change whether payments are observed by other buyers or not.
- Payments are higher if they are made public and if buyers are together with acquainted others.
- The effect is additive as there is no positive interaction effect.

Published:
Journal of Behavioral and Experimental Economics, Vol. 90, February 2021, 101631.
DOI: https://doi.org/10.1016/j.socec.2020.101631

A.4 Gender Bias in the Evaluation of Teaching Materials

(coauthored with Holger A. Rau, Stefan T. Trautmann, and Christian König-Kersting)

Abstract

Gender differences in university teaching evaluations are well established, showing less favorable assessments of female instructors. It has also been shown that these differences cannot be linked to differences in students' course performance, which would justify differences in evaluations. The less favorable assessments are thus either due to differences in aspects that do not affect student performance, but do affect their class experience (e.g., likability of voice tone), or due to evaluation biases unrelated to any actual differences in class experience. We find support for the latter mechanism when any differences between instructors are excluded by having respondents judge identical teaching materials prepared by either a male or a female instructor. In two studies, we find that female instructors receive worse ratings than male instructors from male respondents. In one study, we also find that female instructors receive higher ratings from female raters. Gender bias vanishes for non-academic subjects in our data.

Keywords: Gender equality, Discrimination, Teaching evaluations, Higher education, Experiment

JEL Classifications: C91, D03, L11

Published:

Frontiers in Psychology, Vol. 11, May 2020, Article 1074.

DOI: https://doi.org/10.3389/fpsyg.2020.01074

A.5 Delayed Risks in Individual and Social Decisions

(coauthored with Holger A. Rau and Stefan T. Trautmann)

Abstract

This paper experimentally studies temporal-distance effects on decisions under (social) uncertainty. We elicit subjects' WTPs for risky and ambiguous lotteries with the same expected value. The lotteries differ in the combination of (high vs. low) probabilities with (low vs. high) payoffs. We vary the timing (immediately vs. in two weeks) when the uncertainty is resolved. Subjects always have lower WTPs for longshots than for safer lotteries. Under ambiguity, this gap decreases in the temporal-distance treatment. We find no temporal-distance effect for risky lotteries. Moreover, we find that subjects are always ambiguity averse, which becomes less pronounced when low probability lotteries are processed in the future. In a trust game, we study temporal-distance effects on social risks. Time distance lowers trusting behavior as trustors correctly anticipate that reciprocity is lower when trustees decide in the future.

Keywords: Time distance, construal level theory, ambiguity, individual risks, social risks

JEL Classifications: C91, D03, D81

1. Introduction

Delayed resolution of risks plays a crucial role in economics and management, as the returns of many investments materialize only with some temporal delay. Examples involve financial investments, taking out insurance, research and development strategies, provision of loans, or the execution of a joint project. While the first three examples refer to decision making contexts that are typically less affected by the actions of another person (*individual risk*), the last two examples represent strategic contexts where the uncertainty involved derives from another person's willingness to cooperate (*social risk*).

Construal Level Theory (CLT; Liberman and Trope, 1998; Trope and Liberman, 2000, 2003, 2010) offers a helpful psychological framework for the understanding of temporal delay effects in economic decisions. CLT proposes that individuals make a trade-off between an outcome's *desirability* and its *feasibility*, and weight the outcome *desirability* relatively lower than the outcome *feasibility* in the near future versus the more distant future.[1] For the case of lotteries, it has been shown that *desirability* depends on the level of the outcome and *feasibility* is represented by the level of the winning probability (Sagristano, Trope, and Liberman, 2002; but see Trautmann, 2019). CLT thus implies that in the near future individuals have an increased focus on the winning probability of a lottery, whereas its outcome is less important. By contrast, in the distant future they focus more on the outcome of a lottery, whereas its winning probability is less important. Thus, lotteries characterized by a low (high) winning probability and a high (low) payoff become more (less) attractive when they materialize in the future (Sagristano, Trope, and Liberman, 2002).

In the economics literature, delayed resolution has typically been studied for *known* risks, i.e., in settings in which individuals know the probability distributions of risky outcomes. It has been found that in this case, individuals tend to become more risk tolerant when lotteries materialize with temporal delay (e.g., Noussair and Wu, 2006; Abdellaoui, Diecidue, and Öncüler, 2011). However, most uncertainties outside the experimental lab are characterized by *unknown* (*ambiguous*) risks, i.e., by situations in which individuals do not know the probability distributions of the prospects (Ellsberg, 1961; Trautmann and van de Kuilen, 2015). From a theoretical point of view, CLT should also apply to temporal delay effects for *ambiguous* risk. Moreover, ambiguity might affect the perceived feasibility of the uncertain outcome. That is,

[1] The reason is that the *desirability* and the *feasibility* of an outcome are mentally represented on different construal levels. Higher-level construals, such as outcome *desirability*, are enhanced over delay, while lower-level construals, such as outcome *feasibility*, are discounted over delay (Liberman and Trope, 1998; Sagristano, Trope, and Liberman, 2002; Trope and Liberman, 2003).

decision-makers may perceive ambiguous lotteries as less feasible than risky lotteries, as the probability distribution is unknown. However, there is little evidence on temporal delay effects for ambiguous lotteries versus risky lotteries yet.[2] Therefore, the first goal of this paper is to analyze time delay effects in risky versus ambiguous lotteries.

In many practical settings, uncertain outcomes often also depend on the actions of other decision makers. A decision maker faces *social risk*. These situations therefore represent an important domain to study temporal distance effects, as it is likely that these outcomes are only revealed after some temporal delay. Consider an investment in a joint project. This situation is characterized by trust, as an investor faces the (social) risk that the trustee exploits his trust and does not share the revenue with him in full. Thus, when decision makers face *social risk*, not only uncertainty attitudes, but also social preferences may be affected by temporal delay. Little is known yet about the effects of temporal delay in these situations. Thus, we extend our analysis to the domain of social risk.[3,4]

In this paper, we aim to get a comprehensive understanding of the impact of time delay in decisions under individual risk, both known and unknown, and social risk. In a between-subjects design, we identify the causal effects of time delay by varying the point in time (immediately vs. two weeks later) when the outcomes of the lotteries are materialized. Our design allows a direct test of the widely-cited CLT predictions. In our main treatment (*Late*), subjects state their willingness to pay (WTP) for lotteries, which are resolved two weeks later in the distant future. By contrast, subjects in the control treatment (*Early*) state their WTPs for the same lotteries, which are resolved immediately. Following CLT, we test whether low-probability high-payoff lotteries (low feasibility and high desirability) are perceived as more favorable when they are resolved in the distant future, and vice versa for high- probability low-payoff

[2]Onay, La-Orunal and Öncüler (2013) and Liu and Öncüler (2017) are exceptions. The papers find that temporally delaying the resolution of lotteries may attenuate ambiguity aversion. However, a problem in the experimental design of these papers for our research question is that they not only varied the timing of the resolution of the lotteries, but also the timing of subjects' payments. As a consequence, it is possible that the change in subjects' behavior may also be affected by subjects' time preferences, not uncertainty attitude.

[3]Ederer and Schneider (2020) find in a binary trust game with background risk (i.e., the amount sent back by the trustee is lost with a non-zero probability) that temporally delaying the decision of the trustee did not affect the frequency of trust. We aim to observe changes in the degree of trust in a deterministic setting; we also use a more fine-grained, non-binary setting.

[4]We do not consider the social aspect entering ambiguous decisions in the form of group decision making (Charness et al., 2013; Keck et al., 2014). Although the interaction of social and temporal effect is of much interest for practical applications, the methodological complications of group decision making imply that these effects may better studied independently, in a design tailored to group decision research.

lotteries. We also test the predictions of CLT for ambiguous lotteries, studying behavior for ambiguous events, as well as studying ambiguity attitudes in the delayed-resolution context. The idea is that ambiguous lotteries may be perceived as less feasible, since the probability distribution is not known. We test whether ambiguous lotteries become relatively more attractive compared to risky lotteries in the distant future, as feasibility is less of a concern.

We complement the analysis of individual delayed risks with a situation where trustors in a trust game are confronted with social risk. In the *Late* treatment, trustees are informed about the amount invested by trustors, but they do not take their decisions until two weeks have passed. In the *Early* treatment trustees are informed and decide immediately. Importantly, in all conditions and for all tasks, actual payments will always occur in the second meeting.

Our data support the CLT prediction that low-probability event lotteries become more attractive than high-probability event lotteries for future resolution for ambiguity, and for some specifications for known risks. Focusing on ambiguity attitudes, we find little effect overall. Finally, our analysis of social risk shows that delayed decisions lead to lower trustworthiness, which is correctly anticipated by trustors qualitatively, but not quantitatively: the effect is larger than expected by trustors. Beliefs strongly map into trusting behavior, but the small belief effect triggered by Late resolution induces only a small and insignificant effect on trust.

In sum, our results suggest that some temporal delay effects on individual and social uncertainty can be observed, but that these effects are potentially small. That is, delayed resolution does not qualitatively challenge typical immediate-resolution results, despite some significant changes in behavior. Note that virtually all experimental measurements of individual or social risk are conducted in immediate-resolution designs ("early"), where risk is resolved and subjects are paid directly after the experiment. In contrast, many situations outside the lab involve delayed resolution and payment. Some studies building on CLT have suggested that temporal delay effects might be substantial, which would harm the external validity of immediate resolution-designs for outside-the-lab decision making. Our data, while supporting the presence of temporal effects, do not suggest that substantial threats to external validity derive from differences in the resolution of uncertainty.

2. Experimental Design

We test the effects of temporal distance on risk-taking behavior for individual risks and social risks. In a lottery task, we elicit subjects' valuations of individual risks by asking subjects to state their willingness to pay (WTP) for lotteries with known and

unknown risks. The lotteries are characterized by the same expected value, but they differ in the winning probability (high vs. low). In a trust game (Berg, Dickhaut, and McCabe, 1995), we measure subjects' attitudes towards social risks by studying their trust (or investment) levels in the strategic game.

We manipulate temporal distance with two between-subject time conditions (*Early* treatment and *Late* treatment) to which subjects are randomly assigned. Both conditions consist of two experimental meetings each. In *Early*, all uncertain outcomes of individual and social risks are resolved at the end of the first experimental meeting. By contrast, in *Late*, uncertain outcomes are resolved at the beginning of the second experimental meeting two weeks later. All decision-dependent payments are obtained in the second, later meeting. In what follows, we explain the lottery tasks and the trust game in both time conditions in detail.

2.1 Lottery task

We measure subjects' WTPs for four different lotteries, which are presented in two subsequent subsets of two lotteries each. In each set, subjects are confronted with a risky lottery and an ambiguous lottery, both of which have the same (underlying) objective winning probability and the same monetary payoff. In the first set, subjects have to state their WTPs for playing a risky lottery and an ambiguous lottery with a low payoff of €10 and a high objective winning probability of 0.8. In the second set, subjects have to state their WTPs for playing a risky lottery and an ambiguous lottery with a high payoff of €40 and a low objective winning probability of 0.2. All four lotteries have the same expected value of €8. In the experiment, we vary the order of the two lottery sets for half of the subjects to control for order effects. The monetary payoff of a lottery is determined by a random draw from one of two boxes, each of which is filled with 100 colored chips. For the risky lotteries the random draw is conducted out of the risky box, i.e., a box where the composition of colored chips is known. The risky box contains colored chips of ten different colors, each of which exists exactly ten times. By contrast, for the ambiguous lotteries the random draw is conducted out of the ambiguous box, i.e., a box where the composition of colored chips is unknown. The ambiguous box may contain chips of the same ten colors, but the exact quantity of each color is unknown. Table 1 gives an overview of the four lotteries and the two lottery sets presented to subjects in order 1.

Lotteries were determined as follows. In all sessions, before the decision tasks were explained or started, subjects were first asked to select a set of colors for later use (the "winning" colors for the later random draws). On the computer screen, we presented subjects subsequently with two sets of ten colors. Each set consisted of the

Table 1: Risky and ambiguous lotteries conditional on the representation set.

lottery set 1				
box	**lottery type**	**obj. winning prob.**	**payoff**	**lottery characteristics**
known composition	risky	0.8	€10	high prob. & low payoff
unknown composition	ambiguous	0.8	€10	high prob. & low payoff
lottery set 2				
box	**lottery type**	**obj. winning prob.**	**payoff**	**lottery characteristics**
known composition	risky	0.2	€40	low prob. & high payoff
unknown composition	ambiguous	0.2	€40	low prob. & high payoff

same ten colors (black, red, blue, green, yellow, brown, purple, grey, cyan, white). In the first set we asked subjects to select eight of the ten colors. In the second set they had to select two out of the ten colors. We used subjects' selections in the two sets as a criterion to determine whether a lottery would pay out in the high winning probability (0.8) and low winning probability (0.2) cases, respectively. For instance, if a subject has selected red and black in set two, this means that a lottery of this set would win, if either a red or a black poker chip was drawn. Importantly, at the time of selecting the colors, subjects were not told the purpose of the selection process, nor did they receive any instructions for the experiment yet.

After selecting the two subsets of colors, subjects were presented with the two boxes. For this purpose, we asked them to come to the reception room of the lab. In this room, the two boxes were placed on the table.[5] We explained that each box contained exactly 100 colored chips. We also told them that box 1 (known composition) exactly contained ten different colors, which each contained exactly ten times, and that box 2 (unknown composition) exactly contained 100 colored chips, but the composition of the ten colors was unknown. We informed the subjects that the chips could be of any of the ten colors as in box 1, in an unknown composition. Afterwards, subjects returned to their cubicles in the lab. We handed out paper instructions, explaining the decision-making process for evaluating the four lotteries.

2.1.1 Elicitation of WTPs

We applied a Becker-DeGroot-Marschak (1964) mechanism to elicit subjects' WTPs for the lotteries in an incentive-compatible way. In each set, subjects had to state their WTP for the risky lottery and for the ambiguous lottery. Subjects were endowed with €10 to pay for the lotteries, and they could choose any amount between €0 and €10 in increments of €0.01, respectively. After subjects completed set one, they were presented with set two.

[5]The boxes were not yet placed at the table when subjects arrived at the laboratory.

Once subjects had specified their WTPs for all lotteries, a computerized random draw determined which of the four lotteries was relevant for a subject. Then, the computer determined a random price between €0 and €10 in increments of €0.01 for the relevant lottery. If a subject's WTP was at least as high as the random price, the subject bought the relevant lottery and paid the price determined by the computer. In this case, the random price was subtracted from the initial endowment of €10, with the remainder kept by the subject. If a subject's WTP was below the random price, the subject did not buy the relevant lottery and kept her initial endowment of €10.

In a next step, subjects were informed on the computer screen whether they bought a lottery and if so, which box would be relevant for the random draw. Subjects knew that in the case they have bought a lottery, they would have to draw a chip out of the relevant box to determine whether the lottery they bought would pay out. The random draw was either processed with box 1 or box 2, depending on whether a risky or an ambiguous lottery was chosen by the computerized random draw. If the drawn color matched one of the colors that was initially selected by the subject in the relevant set, the lottery paid out. Otherwise, the lottery did not pay, and the subject received a payoff of €0 from the lottery.

2.1.2 Processing of the lottery outcomes in the *Early* vs. *Late* treatment

The procedure described above was identical in the *Early* and *Late* treatment. In both conditions, subjects were informed in the first experimental meeting on whether or not they bought a lottery. The crucial (and only) difference between *Early* and *Late* was the timing of the random draw of a chip from the selected box. In *Early*, subjects had to draw a chip out of the relevant box at the very end of the first experimental meeting. In *Late*, subjects had to draw a chip out of the relevant box at the end of the second experimental meeting two weeks later. In both conditions, payments were made in the second meeting.

2.1.3 Guessing of the stated WTPs during the second meeting

We also controlled for whether subjects in the second experimental meeting (which took place two weeks after the first meeting) still remembered their stated WTPs for the four lotteries (potentially indicating a clear preference). Subjects in the second experimental meeting were again presented with the same four lotteries of the first experimental meeting. They could earn €0.50 for each correct guess of their own WTPs indicated during the first meeting. Second meetings in both treatments started with this guessing task.

2.2 Trust game

2.2.1 Decisions

In the next stage, we measured the effects of temporal distance on subjects' attitudes towards social risk in a trust game. The trust game was conducted in the first experimental meeting before subjects were informed on whether or not they bought a lottery. At this point, we handed out new instructions, explaining the trust-game stage to the subjects. In the game, we apply the strategy method to elicit first-mover (trustor) decisions, i.e., all subjects decided in the role of a trustor. For this purpose, they were told that they have an endowment of €6, which could be sent in increments of €0.50 to the second mover (i.e., the trustee). Subjects were informed that any positive investment is tripled and sent to the trustee. Subjects knew that after deciding as a trustor, half of them are randomly assigned the role of a trustee. It was made clear that each trustee is paired to an anonymous other subject which was assigned the role of a trustor. The trustees also received an endowment of €6. Trustees could return any amount (in increments of €0.50) of the tripled amount that they received from the trustor. Trustees could not send their endowment to the trustor.

We also elicited trustors' beliefs about trustees' returns for each of six possible full integer investments (€1, €2,...,€6). Subjects gave their unincentivized belief by indicating the expected return- amount by open entry in the interval between €0 and 3 times the transferred investment. The inputs could be made in increments of €0.50. We collected these data right after trustors made their decisions. At this time, they did not know their role in the game yet.

2.2.2 Resolution of the trustee' decisions: *Early* vs. *Late* treatment

The treatments differ in the timing when trustees make their decisions and when trustors are informed on the outcome of the *social risk*. Importantly, in both treatments, trustees observed the decision of the trustor in the first experimental session. However, only in *Early*, trustees made their decisions in the first meeting. By contrast, in *Late*, trustees made their decisions two weeks later in the second meeting. In the *Late* treatment, trustees decided right after the WTP-guessing task. Once the trust game was finished, the experiment continued with the resolution of the random draws of the lottery-valuation task.

2.3 Experimental procedures

2.3.1 Timing of the experimental stages

In both treatments, each session consisted of two experimental meetings with the same subjects. That is, each session was split up to two experimental meetings on different days. The second meeting took place two weeks later on the same day of the week and at the same starting time as the first meeting. We recruited subjects with ORSEE (Greiner, 2015) and the experiment was programmed with z-Tree (Fischbacher, 2007). We informed subjects that in order to participate in the experiment, they would have to sign in for both meetings on both days. They were also informed that they would only receive the compensation of the experiment, if they participated in both meetings.[6] In this case, they will receive the total payments for the decisions made in both meetings after the end of the second meeting. In the *Late* treatment, it is by design not possible to pay subjects in the first meeting, since we do not resolve the lottery outcomes until two weeks later. To rule out any wealth effects or time-discounting effects, we also delayed the payment of subjects' final earnings in the *Early* treatment. The only payment subjects received for participating in the first meeting was a show-up fee of €5. The show up was paid to them after the end of the first meeting in both treatments.

Then subjects took part in the lottery task and completed the trust game. In the *Early* treatment, we resolved the lotteries at the end of the first meeting. Subjects were informed on the computer screen on the relevant lottery and whether they purchased it. Then, we came to each cubicle and checked the computer screen whether they had bought a lottery, and if so, brought the relevant box to their desk. Next, subjects could draw a colored chip out of the relevant box. They won the lottery if the chip was of one of their selected colors. We checked this on the computer screen and typed in the outcome of the lottery in subjects' computer terminal.[7] Next, subjects were handed out envelopes, which contained a sheet of paper. They were told that the computer will allocate a unique id to them, which will be shown on the computer screen. The purpose of this was to store the data on subjects' choices and earnings from the first meeting such that this information could be used in the second meeting to finish the experiment and pay subjects. They were told that they should write down their id and they had to put this sheet of paper in the envelope. Finally, subjects completed a short computerized post-experimental questionnaire,

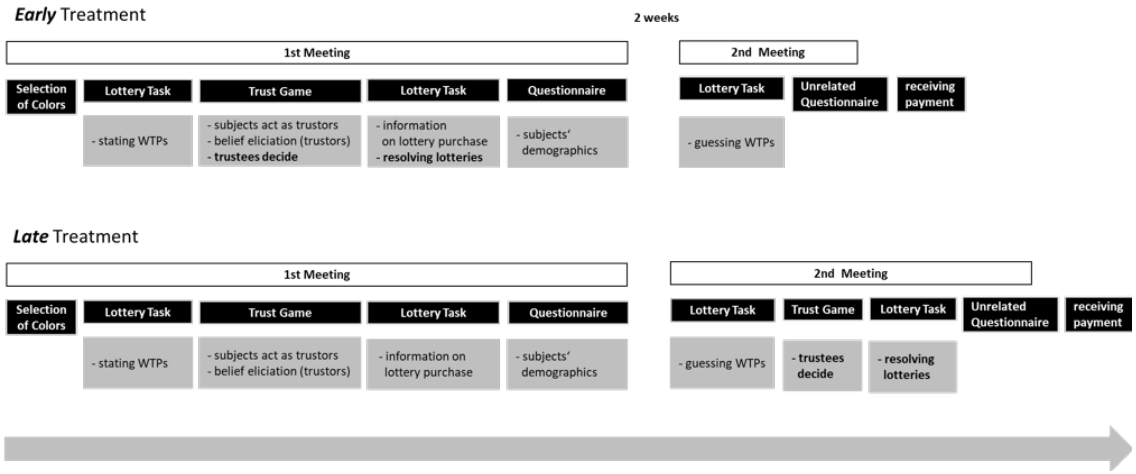
[6] At the beginning of the first experimental meeting, subjects had to sign a form to give their consent to the rule that they will not get their final payments, if they do not show up at the second meeting.

[7] The input was protected by a code which was only known to the experimenter.

where we asked them for their gender, their age, their field of study, and the level of study (bachelor vs. master). After the first meeting was finished subjects were given the €5 show-up fee when they brought the envelope to the experimenter. We kept the envelopes with their names on it until the second meeting took place, when they would open it for the code.

In the *Late* treatment, we did not resolve the lotteries at the end of the first meeting. Here, subjects were only informed about the selected lottery, the drawn random price, and whether they bought the lottery. At the beginning of the second meeting, subjects received the envelope containing their id. They were told that they will receive another €5 for participation in this meeting (this was the same in both treatments). Subjects typed in their id on the computer terminal, and after identification by the computer, we started to elicit subjects' guesses of their WTPs of the first meeting (see Section 2.1.3), followed by the resolution of the risk tasks. The computer reminded them of their role in the trust game, which started two weeks ago. Furthermore, the trustees were informed about the amount that was sent to them by their matched trustor, and they made their decisions. Next, the lotteries were resolved. Afterwards, subjects completed a questionnaire, which was used for an unrelated study (Özgümüs et al., 2020). Then, subjects were informed on the computer screen on their total payments from the first and second experimental meetings. Finally, they were paid out. Figure 1 summarizes the timing of all experimental stages in both experimental meetings of *Early* and *Late*. In the figure, we emphasize any treatment differences in bold font.

Figure 1: overview of all experimental stages in the first and second meetings of our treatments.



2.3.2 Procedural details

The experiment had seven sessions of each treatment. All sessions consisted of two experimental meetings. In total 264 subjects participated, i.e., 132 subjects (63 male, 69 female) in *Early* and 132 subjects (61 male, 71 female) in *Late*. The attrition rate was low, with 249 of the 264 subjects again showing up for the second meeting. Note that all relevant data except trustworthiness were already collected during the first meeting. The first meetings took on average 70 minutes in Early and 60 minutes in *Late*. The second meetings took on average 30 minutes in *Early* and 45 minutes in Late. On average, subjects earned €30.38 in *Early* and €28.82 in *Late*. The study was preregistered on aspredicted.org under the number #18211: https://aspredicted.org/blind.php?x=yc2pt8.

3. Hypotheses

3.1 Individual risk

In this section, we derive our hypotheses for individual risk taking. We build our hypotheses on Construal-Level Theory (CLT, Liberman and Trope, 1998; Trope and Liberman, 2000; 2003; 2010), the leading psychological theory to explain behavior depending on the temporal framing. Applying CLT and following Sagristano, Trope, and Liberman (2002), we predict that temporally distancing the resolution of lotteries increases the preference for low-probability lotteries with high payoffs relative to high-probability lotteries with low payoffs. The reason is that people put a lower weight on the *feasibility* of the lotteries (which is represented by the level of the winning probability) and a higher weight on the *desirability* of the lotteries (which is represented by the level of the possible payoff), for more distant resolution. In our setting, this should translate to a smaller within-subject difference *WTP (high-probability lottery)–WTP (low-probability lottery)* when lotteries are resolved in the future.

H1a: *The difference between WTPs of the risky lotteries with high and low probabilities should be smaller, if the lotteries are processed in the Late treatment compared to Early treatment.*

Translating this reasoning to the domain of ambiguity, we predict that temporally spacing the resolution of ambiguous outcomes should lead to a similar effect. That is, the lower-level construal *feasibility* should be discounted, whereas the higher-level construal *desirability* should be enhanced in the future. Translated to our

ambiguous lotteries, we expect that in the distant future, subjects should show an increased preference for the ambiguous lottery with a low winning likelihood and a high payoff, as they are attracted by the desirability of the high payoff. This should thus again lead to a smaller within-subject difference *WTP (high-likelihood lottery)–WTP (low-likelihood lottery)* when lotteries are resolved in the future.

H1b: *The difference between WTPs of the ambiguous lotteries with high and low winning likelihood should be larger, if the lotteries are processed in the Late treatment compared to Early treatment.*

Regarding ambiguity attitudes, we assume that the unknown probability distribution of ambiguous lotteries leads to a lower perceived feasibility as compared to risky lotteries. Individuals are typically ambiguity averse for high-probability lotteries with low payoff and ambiguity seeking for low-probability lotteries with high payoffs (see Trautmann and van de Kuilen, 2015). This would be our prediction in *Early*, which is the typical design feature in most experimental studies (because payments are obtained immediately). If individuals' valuations of risky and ambiguous lotteries are affected by temporal distance (see H1a, H1b), this may have consequences for individuals' ambiguity attitudes, too. First, if there are different effects for risk and ambiguity regarding low- versus high-likelihood lotteries, this affects the ambiguity attitude as defined by the difference in valuations for risky and ambiguous lotteries with the same underlying probability. Second, the lower *feasibility* of ambiguous lotteries may become less relevant for distant resolution. We thus predict that for high likelihood lotteries, ambiguity aversion should be attenuated when lotteries are resolved in the distant future. Conversely, given our expectation that individuals are ambiguity seeking for low-probability lotteries with high payoffs, the lower focus on the probability dimension in *Late* should lead to a weaker relative preference for ambiguous lotteries in the low-probability case.

H2a: *For high likelihood lotteries, the difference in WTPs of the risky and ambiguous lotteries should be smaller, if the lotteries are processed in the Late treatment compared to Early treatment (less ambiguity aversion).*

H2b: *For low likelihood lotteries, the difference in WTPs of the risky and ambiguous lotteries should be larger (less negative), if the lotteries are processed in the Late treatment compared to the Early treatment (less ambiguity seeking).*

3.2 Social risk

For social risk we do not derive hypotheses, as there are two opposing theories how fist movers may react to time distance. On the one hand, trustors decide under uncertainty and face ambiguity regarding second movers' likelihood of behaving trustworthy. Assuming that first movers face an ambiguous gamble when investing, CLT would predict that trustors trust more in Late compared to Early, due to the lower focus on the probability dimension. On the other hand, there is evidence that time distance may induce a more selfish behavior of trustees. Neo et al. (2013) argue that reciprocity may be attenuated in the future, as affective states are subject to adaption (Frederick and Loewenstein, 1999). This effect may be pronounced, as trustees may behave less reciprocal in the distant future due to an increased feeling of loss aversion. The idea is that trustees perceive a loss when returning money to the trustor, as they believe the received money is their property. Research suggests that this valuation may be amplified over time (Strahilevitz and Loewenstein, 1998). Finally, Neo et al. (2013) argue that time delays should lead to more System-2 processing, which is responsible for less intuitive, but more controlled and slow decisions. Therefore, after time delays, subjects should behave more rational, as hey have more time for System 2 processing.[8] Thus, social preferences are attenuated, and the trustees behave more selfishly. As a consequence, trustors should lower their investments in *Late* compared to *Early*, when they anticipate the attenuated reciprocity of trustees in the distant future. In sum, it is not clear which of the two effects will dominate. Therefore, we do not formulate a hypothesis for social risk. In the results section we test, which of the two effects dominates.

4. Results

4.1 Individual risk

Table 2 presents an overview of the mean WTPs (in Euro) for the risky and ambiguous lotteries, and for high and low likelihood events. The table also reports the difference in the WTPs between the high probability and the low probability lotteries, as well as between the risky and ambiguous lotteries for a given probability dimension (0.8 vs. 0.2). Results are shown by treatments. The data encompass n = 264 observations (n = 132 in each of the two treatments). We first present raw comparisons using non- parametric tests, which test for differences in subject's WTPs within the treatments. We then turn to regression analyses to test for treatment

[8]This might explain similar findings observed in other situations of social risk such as the dictator game or ultimatum game (Kovartik, 2009; Grimm and Mengel, 2011).

effects in the difference in WTP-difference between *Early* and *Late*.

Table 2: Subjects' WTPs (in Euro) for the risky and ambiguous lotteries in Early and Late.

	***Early* Treatment**			***Late* Treatment**		
	Prob. 0.8	Prob. 0.2	Difference	Prob. 0.8	Prob. 0.2	Difference
Risky Lottery	6.32 (2.63)	5.06 (2.65)	**1.26 (3.12)**	5.49 (2.72)	4.80 (3.05)	**1.14 (3.69)**
Ambig. Lottery	4.79 (2.87)	3.68 (2.58)	**1.11 (2.63)**	4.58 (2.98)	3.84 (3.00)	**0.74 (3.74)**
Difference	**1.52 (2.46)**	**1.38 (2.40)**		**1.36 (2.50)**	**0.96 (2.33)**	

Notes: Bold print indicates significant differences at the 1% level, Wilcoxon signed-rank tests. Standard deviations in parentheses.

For the risky lotteries, subjects express a higher WTP in the high probability tasks as compared to the low probability tasks (Wilcoxon signed-rank tests, $p < 0.001$, in both treatments), and no evidence for risk-seeking (which would imply WTP>€8) for low probability lotteries. The difference in the WTPs between the high vs. low probability lotteries is moderately lower in the *Late* condition (1.14) as compared to the *Early* condition (1.26). For the ambiguous lotteries, we also find that in both treatments the WTP is significantly higher in high probability tasks than in the low probability tasks (Wilcoxon signed-rank tests, $p < 0.001$, in both treatments). The difference in the WTPs between the high and low probability lotteries is lower in *Late* (0.74) as compared to *Early* (1.11). Subjects are ambiguity averse in both tasks of both treatments, i.e., the WTPs of risky lotteries are significantly higher than the WTPs of ambiguous lotteries for both high and low probability (Wilcoxon signed-rank tests, $p < 0.001$). Thus, for low-probability lotteries, our data show no evidence for ambiguity-seeking (which would imply WTP-ambiguous > WTP-risky). The WTP difference between risky and ambiguous lotteries is slightly lower in *Late* (1.36 and 0.96) than in *Early* (1.52 and 1.38). A remarkable feature of our data is the absence of risk and ambiguity seeking behavior for low likelihood events, in contrast to the fourfold patterns often reported in the literature (e.g., Trautmann and van den Kuilen, 2015; Kocher et al., 2018; Li et al., 2020). The strongly and universally risk and ambiguity averse behavior might be explained by the WTP approach. WTP emphasizes reference points and framing of buying price as loss (Trautmann et al., 2011); this might have led subjects to bid lower values for the low-probability lotteries compared to choice-list or WTA-based studies (which measure certainty equivalents). Since we focus on differences between conditions, the absolute price level should not be crucial for our research question.

We test for treatment effects in the WTP differences (H1a and H1b) with random

effects OLS regressions on the WTP differences in subjects' valuation of high vs. low probability lotteries. Since all subjects decide on the WTPs of low and high probability lotteries in the risky and in the ambiguous domain, we can compute two WTP differences for each subject. We consider this panel structure in our regression analyses. In model (1) we include the treatment dummy of intertest (*Late Condition*), a dummy for risky lotteries (*Risky Lottery*), and the interaction term *Late Condition x Risky Lottery* that measures the differential effect of Late in the risky-lottery conditions. Finally, we add a dummy (*no-show*), which is positive for the 15 subjects, who did not show up for the second experimental session. In models (2) and (3) we include control variables for gender, age, whether subjects are economics students and for the presentation order (1 or 2) of the lotteries. We also add a variable that controls for the size of the experimental session.[9] Finally, in model (3) we include the differences between subjects' guessed WTPs (stated in the second experimental session) and their real WTPs (stated in the first experimental session), as a measure of the robustness of their bids (assuming noisy WTP are remembered less precisely). The regressions are clustered at session level.

Table 3: Random effects OLS regressions on the valuation of high prob. vs. low prob. lotteries.

	WTP Difference (high prob. - low prob.)		
	(1)	(2)	(3)
Late Condition	-0.38 (0.22)*	-0.40 (0.20)**	-0.57 (0.21)***
Risky Lottery	0.15 (0.12)	0.15 (0.12)	0.19 (0.12)*
Late Condition x Risky Lottery	0.25 (0.21)	0.25 (0.21)	0.13 (0.23)
No-Show	0.49 (1.04)	0.34 (1.05)	
Guess WTP risk high			0.01 (0.02)
Guess WTP ambig high			0.02 (0.03)
Guess WTP risk low			0.02 (0.03)
Guess WTP ambig low			-0.03 (0.04)
Controls	No	Yes	Yes
N WTPs (subjects)	528 (264)	528 (264)	498 (249)

Notes: Report random effects panel regression results, robust standard errors in parentheses. Standard errors are clustered on 14 sessions. Controls are order of the presentation of the lotteries, session size, female, age and a dummy, controlling for economics students. *,**,*** indicate significance at the 10%, 5%, and 1% level, respectively. Sample sizes vary across columns due to some missing values.

[9] As the size of our sessions varied between 10 and 24 subjects, we control for potential session effects, caused by the number of subjects participating in a session.

Model (1) shows a marginally significant difference between WTP-differences in Early and Late. In models (2) and (3), the effect is larger and significant. The observed effect is in line with the predictions of CLT. That is, the WTP difference between high and low probability lotteries becomes significantly smaller when lotteries are processed two weeks later. F-tests show that *Late Condition* and *Late Condition x Risky Lottery* are jointly significantly different from zero in model (3) (p = 0.030), but not in model (1) (p = 0.214) and in model (2) (p = 0.121). That is, the effect is mostly present for ambiguous lotteries, and less clear for risky lotteries. This is consistent with the raw averages shown in Table 2. That is, we find clear support for H2a, less so for H1a.

Next, we test for time-distance effects on subjects' ambiguity attitudes (H2a, H2b). Table 4 presents random effects OLS regressions on the WTP differences between risky and ambiguous lotteries. The WTP difference can be interpreted as a subject's ambiguity attitude, i.e., subjects with a positive (negative) difference can be interpreted as ambiguity averse (ambiguity seeking), whereas subjects with no WTP difference show ambiguity neutrality. We present similar regression models as in Table 3. The only difference is that we now incorporate an indicator variable (*Low prob. Lottery*), which is positive (zero) for low (high) probability lotteries, instead of the *Risky Lottery* indicator. Moreover, we add the interaction between this variable and the treatment dummy (*Late Condition x Low prob.*).

The treatment dummy is not significant in any of the models. Thus, we find no treatment effects for ambiguity attitudes when lotteries are processed in the future. This holds for both low and high probability lotteries. Thus, we reject hypotheses 2a and 2b. The moderately significant negative coefficient of *Low prob. Lottery* suggests in model (3) that ambiguity aversion is generally less pronounced for low probability lotteries, which is in line with the literature (e.g., Trautmann and van den Kuilen, 2015; Kocher et al., 2018).

Table 4: Random effects OLS regressions on the ambiguity attitude for high prob. and low prob. lotteries

	WTP Difference (risky - ambig. lotteries)		
	(1)	(2)	(3)
Late Condition	-0.19 (0.34)	-0.19 (0.33)	-0.22 (0.28)
Low prob. lottery	-0.15 (0.12)	-0.15 (0.12)	0.20 (0.12)*
Late Condition x Low prob.	-0.25 (0.21)	-0.25 (0.21)	-0.13 (0.23)
No-Show	0.44 (0.73)	0.74 (0.71)	
Guess WTP risk high			0.02 (0.02)
Guess WTP ambig high			-0.01 (0.01)
Guess WTP risk low			0.02 (0.01)
Guess WTP ambig low			-0.04 (0.02)
Controls	No	Yes	Yes
N WTPs (subjects)	528 (264)	528 (264)	498 (249)

Notes: Report random effects panel regression results, robust standard errors in parentheses. Standard errors are clustered on 14 sessions. Controls are order of the presentation of the lotteries, session size, female, age and a dummy, controlling for economics students. * indicates significance at the 10% level. Sample sizes vary across columns due to some missing values

4.2 Social risk

We next turn to the analysis of time-distance effects under social risks. Table 5 presents the data of the trust game. The table summarizes the means of the investment levels of trustors (first movers), of first movers' mean beliefs on the returned amounts, the returned amounts by the trustees (second movers), and the levels of trustees' trustworthiness (return rate). We measured first movers' beliefs on the return rate of the second movers, conditional on different investment levels (€1-€6). Based on this, we calculate the mean beliefs on the returned amount for each treatment. The level of trustees' trustworthiness is the return rate, i.e., the amount which was returned to the first movers divided by their investment. The data are conditioned on the *Early* and the *Late* condition. We have a higher number of observations for first movers, because every player decided in the role of a first mover, but only half of the subjects acted as second movers. For second movers we do not exactly have half the data, as we had to drop subjects who did not show up for the second experimental session.

Table 5: Investment and returns in the trust game

	***Early* Condition**	**N**	***Late* Condition**	**N**
Investment Level of FMs	4.26 (2.01)	132	4.04 (2.13)^{ns}	132
Mean Belief on Returned Amount of FMs	4.78 (1.57)	132	4.31 (1.75)**	132
Returned Amount of SMs	5.89 (4.24)	66	4.44 (4.22)**	56
Received Return Rate of FMs	126.83%	66	92.29%	56

Notes: Comparison between Early and Late treatment: ns = not significant; ** significant at 5% level, Mann-Whitney tests.

The data show that investment levels are approximately five percent lower in the *Late* condition (4.04) than in the *Early* condition (4.26). The difference is not significant though, for a non-parametric test (two-sided Mann-Whitney test, p = 0.334). It can be seen that the mean belief on returned amount is significantly lower in *Late* (4.31) than in *Early* (4.78) (two-sided Mann-Whitney test, p = 0.041). We also find that trustees send back less, when the decision to return is delayed: they return an average amount that is significantly lower by 25 percent in *Late* (4.44) than in *Early* (5.89) (two-sided Mann- Whitney test, p = 0.049). First movers in the *Late* treatment receive a significantly lower return rate (92.29%) than in the *Early* treatment (126.83%) (two-sided Mann-Whitney test, p = 0.022).

Table 6 presents Tobit regressions to test for treatment differences in trust levels. The models focus on the investment levels of first movers. All models incorporate our treatment dummy. The second model adds controls, i.e., *session size*, a variable that controls for the size of our experimental sessions, a *female* dummy, subjects' *age* and a variable, which controls for *economics* students. The third model adds *Mean Belief* (see Table 5). All models are left censored at 0 and right censored at 6. The regressions are clustered at the session level.

Table 6: Tobit regressions on Trustors' investment levels

	Invetsment Level of FMs		
	(1)	(2)	(3)
Late Condition	-0.61 (0.45)	-0.60 (0.48)	-0.19 (0.51)
Mean Belief			0.78 (0.15)***
No-show	2.67 (1.17)**	2.48 (1.17)**	1.86 (1.06)*
Controls	No	Yes	Yes
N	264	264	264

Notes: Columns (1)–(2) report Tobit regression results, robust standard errors in parentheses. Standard errors are clustered on 14 sessions. Controls are session size, female, age and a dummy, controlling for econ students. *,** ,*** indicate significance at the 10%, 5% and 1% level.

The regression models (1) and (2) confirm the findings of the non-parametric tests, i.e., in the *Late* condition, we find an insignificantly lower investment level as compared to *Early*. Model (3) shows that the coefficient of *Mean Belief* is positive and significant, indicating that first movers' beliefs about the other player's reaction have a strong influence on their investment levels. This effect replicates previous findings (Fehr et al., 2020). As shown in Table 5, first movers' beliefs about the amount returned are lower the in *Late* (4.31) than in *Early* (4.78). Thus, they seem to anticipate that reciprocity is attenuated in the distant future (Neo et al., 2013). The small and insignificant drop in trust for *Late* is consistent with a small effect on beliefs transmitted by an effect of 0.78 to investment levels.

Table 7: Tobit regressions on Trustees' returned amount and returned rate of percentage

	Returned Amount of SMs		**Received Return Rate of FMs**	
	(1)	(2)	(3)	(4)
Late Condition	-1.91 (0.87)**	-1.80 (0.80)**	-0.41 (0.15)***	-0.40 (0.12)***
Trust Level	0.58 (0.25)**	0.85 (0.25)**	0.16 (0.05)***	0.16 (0.05)***
Controls	No	Yes	No	Yes
N	122	122	122	122

Notes: Columns (1)–(4) report Tobit regression results, robust standard errors in parentheses. Standard errors are clustered on 14 sessions. Controls are sessionsize, female, age and a dummy, controlling for econ students. **,*** indicate significance at the 5% and 1% level, respectively.

Turning to second movers, Table 7 presents Tobit regressions on trustees' returned amount (models (1) and (2)). These regressions are left censored at 0 and right censored at 18. The table also presents Tobit regressions on the received return rate of FMs (models (3) and (4)), which represents second mover's trustworthiness. These regressions are left censored at 0 and right censored at 3. All models in Table 7 include the same covariates as in the regressions of Table 6. Moreover, we control for the level or trust, because larger investments may elicit higher return rates. Focusing on the returned amounts, it can be seen in model (1) and (2) that the treatment dummy has a significantly negative coefficient. Thus, time distance lowers the amount returned by trustees. Models (3) and (4) confirm that time distance attenuates reciprocal behavior: conditional on the amount invested, trustees return lower amounts. It is noteworthy that the actual reduction due to delay is larger than the one expected by FMs (Table 5); however, this is caused by FM being too optimistic in the *Early* condition.

5. Conclusion

In this paper, we analyzed the effects of the delayed resolution of individual and social uncertainty. Our study is framed in the context of Construal Level Theory (CLT; Liberman and Trope, 1998; Trope and Liberman, 2000, 2003, 2010). We employed an incentivized lab experiment with a sizable sample (n = 264), in which we varied the timing of the resolution of uncertainty (immediately vs. in two weeks). We eliminated time-preference effects by paying all uncertain earnings in the second meeting in both conditions. Extending previous work by Sagristano, Trope, and Liberman (2002) on known risk to ambiguous risks and social risks, we find some evidence for CLT predictions and delay effect. This in itself is remarkable given previous non-replication results for CLT (Trautmann and van de Kuilen, 2012; Trautmann, 2019). However, consistent with papers that question whether reported CLT effect sizes in the psychological literature represent robust phenomena, we find small and sometimes insignificant effects that do not change any qualitative findings (risk aversion, ambiguity aversion, trust and the role of beliefs). One of the strongest effects obtains for a decision under *certainty*, the decision of the trust game second movers. The finding that trustees become more egoistic is in line with the results of Grimm and Mengel (2011). The authors report that delaying the decisions of recipients attenuates social preferences and increases the acceptance of low offers in an ultimatum game.

Our theoretical framework differs from those often applied in economics in two essential features. First, valuation of uncertainty resolution is time-dependent, but not

dynamic (e.g., Kreps and Porteus, 1978, 1979; Ahlbrecht and Weber, 1997). Second, our framework does not shift both the resolution of uncertainty and WTP-payments into the future, and so does not ask individuals to state present estimates of future lottery valuations in order to eliminate time preference effects (e.g., Abdellaoui, Diecidue, and Öncüler, 2011; Pan, Web, and Zank, 2019). Such an approach may lead to potential biases in the desirability of future outcomes, e.g., if individuals anticipate that buying prices are deducted from the initial endowment in the future. We hold constant the date of payments for both WTP and resulting payoffs in the two delay conditions. This enables us to rigorously test time- dependent reversals in both individual and social preferences over the resolution of uncertainty.

In conclusion, we find some support for time distance effects on decisions under uncertainty, which are consistent with the psychological predictions of construal level theory. With the "standard" or "typical" design in experimental decision-making studies employing immediate resolution of risk, and many situations outside the lab involving delayed resolution, external validity may be systematically threatened. However, the effects are small and do not change any qualitative interpretations in our study. In line with previous studies testing the robustness of CLT, we therefore conclude that psychological temporal distance effects are subtle psychological effects that may only have secondary relevance in most practical settings. The external validity of qualitative results of lab experimental studies on decision under uncertainty will not be severely affected if temporal resolution in the lab differs from temporal resolution outside the lab.

References

Abdellaoui, M. Diecidue, E. and Öncüler, A. (2011). Risk Preferences at Different Time Periods: An Experimental Invetsigation. *Management Science* 57, 975-987.

Ahlbrecht, M. and Weber, M. (1997). Preference for Gradual Resolution of Uncertainty. *Theory and Decision* 43, 167-185.

Becker, G. M., Degroot, M. H. and Marschak, J. (1964). Measuring Utility by a Single-Response Sequential Method. *Behavioral Science* 9, 226-232.

Berg, J., Dickhaut, J. and McCabe (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior* 10, 122-142.

Charness, G., Karni, E., and Levin, D. (2013). Ambiguity attitudes and social interactions: An experimental investigation. *Journal of Risk and Uncertainty* 46, 1-25.

Ederer, F. and Schneider, F. (forthcoming). Trust and Promises over Time. *American Economic Journal: Microeconomics*.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics* 75, 643-669.

Fehr, D., Rau, H., Trautmann, S. T. and Xu, Y. (2020). Inequality, Fairness and Social Capital. *European Economic Review* 129, 103566.

Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental economics* 10, 171-178.

Greiner, B. (2015). Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE. *Journal of the Economic Science Association* 1, 114-125.

Grimm, V. and Mengel, F. (2011). Let Me Sleep on it: Delay Reduces Rejection Rates in Ultimatum Games. *Economics Letters* 111, 113-115.

Frederick, S. and Loewenstein, G. (1999). 16 Hedonic Adaptation, in D. Kahneman, E. Diener and N. Schwarz (eds), Well-Being. The Foundations of Hedonic Psychology, NY: Russell Sage, 302-329.

Keck, S., Diecidue, E., and Budescu, D. (2014). Group Decisions under Ambiguity: Convergence to Neutrality. *Journal of Economic Behavior and Organization* 103, 60-71.

Kocher, M. G., Lahno, A. M. and Trautmann, S. T. (2018). Ambiguity Aversion is not Universal. *European Economic Review* 101, 268-283.

Kovartik, J. (2009). Giving Now or Later: Altruism and Discounting. *Economics Letters* 102, 152-154.

Kreps, D. M. and Porteus, E. L. (1978). Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica* 46, 185-200.

Kreps, D. M. and Porteus, E. L. (1979). Temporal von Neuman Morgenstern and Induced Preferences. *Journal of Economic Theory* 33, 199-231.

Li, C., Turmunkh, U., and Wakker, P. P. (2020). Social and Strategic Ambiguity versus Betrayal Aversion. *Games and Economic Behavior* 123, 272-287.

Liberman, N. and Trope, Y. (1998). The Role of Feasibility and Desirability Considerations in Near and Distant Future Decisions: A Test of Temporal Construal Theory. *Journal of Personality and Social Psychology* 75, 5-18.

Liu, Y. and Öncüler, A. (2017). Ambiguity Attitudes over Time. *Journal of Behavioral Decision Making* 30, 80-88.

Neo, W. S., Yu, M., Weber, R. A. and Gonzalez, C. (2013). The Effects of Time Delay in Reciprocity Games. *Journal of Economic Psychology* 34, 20–35.

Noussair, C. and Wu, P. (2006). Risk Tolerance in the Present and the Future: An Experimental Study. *Managerial and Decision Economics* 27, 401-412.

Özgümüs, A., Rau, H. A., Trautmann, S. T. and König-Kersting, C. (2020). Gender Bias in the Evaluation of Teaching Materials. *Frontiers in Psychology* 11, 1074.

Onay, S., La-Ornual, D. and Öncüler, A. (2013). The Effect of Temporal Distance on Attitudes toward Imprecise Probabilities and Imprecise Outcomes. *Journal of Behavioral Decision Making* 26, 362-374.

Pan, J., Webb, C. S. and Zank, H. (2019). Delayed Probabilistic Risk Attitude: A Parametric Approach. *Theory and Decision* 87, 201-232.

Sagristano, M. D., Trope, Y. and Liberman, N. (2002). Time-Dependent Gambling: Odds Now, Money Later. *Journal of Experimental Psychology: General* 131, 364-376.

Strahilevitz, M. A. and Loewenstein, G. (1998). The Effect of Ownership History on the Valuation of Objects. *Journal of Consumer Research* 25, 276-289.

Trautmann, S. T. (2019). Distance from a Distance: The Robustness of Psychological Distance Effects. *Theory and Decision* 87, 1–15.

Trautmann, S. T. and van de Kuilen, G. (2012). Prospect Theory or Construal Level Theory? Diminishing Sensitivity vs. Psychological Distance in Risky Decisions. *Acta Psychologica* 139, 254–260.

Trautmann, S.T. and van de Kuilen, G. (2015). Ambiguity Attitudes, in: G. Keren and G. Wu (eds), Blackwell Handbook of Judgment and Decision Making 1, 89-116.

Trautmann, S. T., Vieider, F. M. and Wakker, P. P. (2011). Preference Reversals for Ambiguity Aversion. *Management Science* 57, 1320-1333.

Trope, Y. and Liberman, N. (2000). Temporal Construal and Time-Dependent Changes in Preference. *Journal of Personality and Social Psychology* 79, 876-889.

Trope, Y. and Liberman, N. (2003). Temporal Construal. *Psychological Review* 110, 403-421.

Trope, Y. and Liberman, N. (2010). Construal-Level Theory of Psychological Distance. *Psychological Review* 117, 440-463.

Ph.D. Program in Economics
Declaration of contribution to each paper of the cumulative dissertation (based on the CRediT taxonomy by Brand et al. (2015)*)

To the five essays of the cumulative dissertation I contributed as follows:

1. **Trust, Vulnerability and Trustworthiness** (coauthored with Claudia Keser): Methodology, Software, Formal Analysis, Investigation, Resources, Data Curation, Writing (both Original Draft and Review & Editing), Visualization.

2. **An Experimental Investigation of Rating-Market Regulation** (coauthored with Claudia Keser, Emmanuel Peterlé, and Martin Schmidt): Conceptualization, Methodology, Software, Formal Analysis, Investigation, Resources, Data Curation, Writing (both Original Draft and Review & Editing), Visualization.

3. **Who is Watching me? Disentangling Audience and Interpersonal Closeness Effects in a Pay-What-You-Want Context** (coauthored with Elisa Hofmann, Michael E. Fiagbenu, Amir M. Tahamtan, and Tobias Regner): Conceptualization, Methodology, Software, Formal Analysis, Writing - Editing & Reviewing.

4. **Gender Bias in the Evaluation of Teaching Materials** (coauthored with Holger A. Rau, Stefan T. Trautmann, and Christian König-Kersting): Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing (both Original Draft and Review & Editing), Visualization.

5. **Delayed Risks in Individual and Social Decisions** (coauthored with Holger A. Rau and Stefan T. Trautmann): Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing (both Original Draft and Review & Editing), Visualization.

Date, Signature

*Brand, A., Allen, L., Altman, M, Hlava, M., and Scott, Jo. (2015). Beyond Authorship: Attribution, Contribution, Collaboration, and Credit. *Learned Publishing* 28, 151-155. DOI: https://doi.org/10.1087/20150211

Ph.D. Program in Economics
Declaration for admission to the doctoral examination

I confirm

1. that the dissertation "Five Essays on Experimental Economics" that I submitted was produced independently without assistance from external parties, and not contrary to high scientific standards and integrity,

2. that I have adhered to the examination regulations, including upholding a high degree of scientific integrity, which includes the strict and proper use of citations so that the inclusion of other ideas in the dissertation are clearly distinguished,

3. that in the process of completing this doctoral thesis, no intermediaries were compensated to assist me neither with the admissions or preparation processes, and in this process,

    - no remuneration or equivalent compensation were provided
    - no services were engaged that may contradict the purpose of producing a doctoral thesis

4. that I have not submitted this dissertation or parts of this dissertation elsewhere.

I am aware that false claims (and the discovery of those false claims now, and in the future) with regards to the declaration for admission to the doctoral examination can lead to the invalidation or revoking of the doctoral degree.

_____

Date, Signature