Psychological mechanisms in strategic interaction under uncertainty

vorgelegt von
M.Sc.
Florian Artinger
aus Berlin

von der Fakultät VII - Wirtschaft & Management
der Technische Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Wirtschaftswissenschaften
- Dr. rer. oec. -

genehmigte Dissertation

Promotionsausschuss:

Vorsitzende: Prof. Dr. Radosveta Ivanova-Stenzel
Berichtende: Prof. Dr. Dorothea Kübler
Berichter: Prof. Dr. Gerd Gigerenzer

Tag der wissenschaftliche Aussprache: 8. Mai 2012

Berlin 2012
D 83

Zusammenfassung

Diese Dissertation umfasst drei eigenständige Essays, die sich mit den psychologischen Mechanismen in strategischer Interaktion beschäftigen. Ökonomische Theorien werden hierbei aus einer empirischen Perspektive untersucht. Das erste Essay ist eine Studie zum Zusammenhang zwischen sozialen Präferenzen und individuellen Unterschieden in den Dispositionen ‚Theory of Mind' und Empathie. Dies wird anhand eines Laborexperiments untersucht. Das zweite Essay beschäftigt sich mit der Rolle von exogenem Risiko in einem sozialen Dilemma in einem Laborexperiment. Im dritten Essay werden Preisstrategien von Gebrauchtwagenhändler und in wie weit diese auf leicht zugängliche Markinformationen zurückgreifen untersucht. Dies wird anhand von Interviews und Daten aus einem Online Gebrauchtwagenmarkt gemacht.

Stichwörter: Begrenzte Rationalität; Soziale Präferenzen; Dispositionen, Individuelle Unterschiede; Soziales Dilemma; Risiko; Entscheidungen aus Erfahrung; Rigide Preise; Preisunterschiede; Suche.

Abstract

This dissertation consists of three self-contained essays concerned with psychological mechanisms in strategic interaction. This is examined from an empirical perspective shedding light on economic theories. The first essay investigates the relationship between social preferences and individual differences in dispositions for theory of mind and empathy using a laboratory experiment. The second essay examines the role of risky outcomes in a social dilemma. The third essay investigates the pricing strategies of used car dealers and to what extent they make use of readily available market information. The data stems from interviews with dealers and data from an online used car market.

Keywords: Bounded Rationality; Social Preferences; Dispositions; Individual Differences; Social Dilemma; Risky Choice; Decisions from Experience; Sticky Prices; Price Dispersion; Search.

Rechtliche Erklärung

Ich erkläre, dass die geltende Promotionsordnung der Technischen Universität Berlin bekannt ist. Die vorliegende Arbeit ist selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe. Die Arbeit wurde noch keiner Prüfungsbehörde in gleicher oder ähnlicher Form vorgelegt.

Diese Dissertation ist aus Forschungsarbeiten mit den zu Beginn des jeweiligen Kapitels genannten Personen entstanden. Das erste Essay wurde als Workingpaper in der der Serie der Friedrich-Schiller-Universität Jena und des Max-Planck-Instituts für Ökonomie veröffentlicht.

Ich bezeuge durch meine Unterschrift, dass meine Angaben über die bei der Abfassung meiner Dissertation benutzten Hilfsmittel, über die mir zuteil gewordene Hilfe sowie über frühere Begutachtungen meiner Dissertation in jeder Hinsicht der Wahrheit entsprechen.

Berlin, den 16. März 2012 Florian Artinger

Contents

Contents

Tables

List of Figures

Figures

Introduction

Psychological mechanisms in strategic interaction

Friedman (1953) proposed that economic theory should only be judged by its accuracy of predictions about observed outcomes and not by the underlying assumptions. The 'as-if' approach models decision making in terms of a perfectly rational profit maximizer. It has been under discussion particularly as decision making is often subject to constraints in knowledge, time, and computational resources (Simon, 1956). This has led to the rise of behavioral economics which aims at incorporating psychological findings and at the same time maintaining rationality as the basic mechanism to model decision making (Berg & Gigerenzer, 2010; Camerer & Loewenstein, 2004).

An alternative concept is the idea of ecological rationality that describes the fit between the decision process and the environment (Gigerenzer, Todd, & the ABC Research Group, 1999; Smith, 2008). A particular illuminating example for ecological rationality in strategic interaction provides the game of chess. Chess is a complex strategic game for which boundedly rational, human, players have proven capable to outsmart brute computational force for a long time (De Groot, 1978). While early computer pioneers such as Newell and Simon in 1957 predicted that a computer would beat the best human chess player within ten years (Simon & Schaeffer, 1992), only in 1997 the computer Deep Blue, with the capacity to calculate 200 million positions per second, defeated the then reigning world champion in chess, Garry Kasparov.

The success of the decision process of grand masters in chess who outperformed computers for many decades relies heavily on pattern recognition built during extensive periods of training. In contrast, the solution put forth by classical game theory which builds on Friedman (1953) propositions, suggests to follow every branch of the game tree to win, loss, or draw. Using backward induction the best strategy is computed. However, any chess player is faced with an estimated 10^{44} possible games for which backwards induction is computationally intractable (Simon & Schaeffer, 1992). This strongly suggests that economic theory can benefit greatly from taking a psychological perspective.

Introduction

This dissertation provides an empirical approach to different psychological mechanisms in strategic interaction. It highlights strengths and weaknesses of these and how they relate to economic theory.

Preferences and dispositions

Stable preferences are a central assumption for economic modeling. In strategic interaction a large amount of evidence, mainly from neutrally framed laboratory games, shows that people are not exclusively self-interested but differ in their social preferences (for a review see Fehr, Schmidt, Kolm, & Ythier, 2006). The first paper in this dissertation[1] examines the link between social preferences and individual differences in the dispositions for theory of mind (ToM) and empathy. These have been found to positively correlate with fairness concerns, which has led a number of authors to suggest that that both potentially underlie theories of social preferences (e.g., Paal & Bereczkei, 2007; Singer & Fehr, 2005; Singer & Lamm, 2009; Takagishi, Kameshima, Schug, Koizumi, & Yamagishi, 2010). A positive correlation between ToM and fairness concerns seems quite surprising as accurate assessment of what another person does facilitates the pursuit of one's own material interest, an assumption that is at the base of game theory.

Employing accuracy of beliefs as a proxy for ToM, we can indeed show that it is applied differently depending on the motivation of a person. Moreover, we do not find any correlation between behavior in two neutrally framed laboratory games on fairness concerns and dispositional measures for ToM and empathy. We also do not find any relation between the dispositional measure for ToM and accuracy of believes. These results shed doubt on the link between social preferences and dispositions.

The concepts of stable dispositions and preferences neglect an important element that Simon (1956) highlights: decision making is frequently the result of an interaction between mind and environment. Dispositions and preferences take on an internalistic view, neglecting the fit with the environment.

---

[1] Artinger, F., Exadaktylos, F., Koppel, H., Sääksvuori, S.. Putting yourself in others' shoes: Are fairness concerns shaped by individual differences in theory of mind and empathy?

Weight-and-add

In a social dilemma people face a situation where it is in the common interest to cooperate but individually rational to free ride. Often, the outcomes might not be given for sure but are subject to a risky environment. Risk has been modeled in Expected Value Theory (EVT) as adding outcomes weighted by their probability. This weight-and-add approach has been maintained in subsequent theories that seek to account for violations of EVT, such as Prospect Theory (PT) (Kahneman & Tversky, 1979). Recent studies find that in a stochastic social dilemma behavior displays loss aversion, a core principle of PT. (Iturbe-Ormaetxe, Ponti, Tomás, & Ubeda, 2011; McCarter, Rockmann, & Northcraft, 2010). A further principle that PT highlights is that people act as-if they overvalue small probabilities and undervalue large probabilities. Yet, this only holds if people decide based on a description of outcomes and probabilities (decision from description). The pattern is inversed if people sample the distribution of outcomes (decisions from experience) (Hertwig, Barron, Weber, & Erev, 2004). The difference has been labeled description-experience (DE) gap and highlights the importance of the way information is presented.

Paper two[2] addresses following two issues: i) how do people respond to differences in the degree of risk in a stochastic social dilemma, i.e., is over- and undervaluing of probabilities as proposed by PT also present in a strategic context? ii) is the DE gap is preserved in a stochastic social dilemma? We thereby compare choices in lotteries to those in stochastic social dilemmas with an equivalent degree of environmental risk. A novel adaptation of PT designed to compare behavior across lotteries and games shows that predictions of PT are not confirmed, neither for lotteries nor games. However, we find that the behavioral pattern in lotteries mimics that in stochastic game. It is then quite surprising that the DE gap holds for lotteries but not for the stochastic games. Process data and subjects self-reported reasons for cooperation suggests that the disappearance of the DE gap in games may result from a decision process that emphasizes the size

[2] Artinger, F., Fleischhut, N., Levati, V., Stevens, J.. Cooperation in a risky environment: Decisions from experience in a stochastic social dilemma.

of the outcomes and expectations about others' behavior over outcome probabilities.

The results suggest that people do not integrate all the available information in such a way that adding a stochastic element to strategic interaction preserves the properties of decisions under risk. Instead, certain elements become more important such as outcomes and expectations about others' behavior whereas other elements loose importance, such as probabilities. In order to understand more complex environments and thus move closer to the real world these results suggest that one needs to gain an understanding which elements of a decision environment dominate under what conditions. This also suggests that lexicographic strategies, where elements of the decision environment are evaluated sequentially instead of integrating them all at once by weighting and adding, might often be a good approximation of human decision making (e.g., Luce, 1956).

Information use

Originating in psychology, two different traditions on heuristics have emerged that both point out that decision makers often rely only on a limited amount of information: heuristics and biases (Tversky & Kahneman, 1974) and fast and frugal heuristics (Gigerenzer et al. 1999). Heuristics and biases emphasize that the decision strategies that people use frequently fall short of a normative standard, usually a perfectly rational profit maximizer of the as-if tradition by Friedman (1953). In contrast, fast and frugal heuristics suggest that decision making is the result of an adaptive process that emerges from the interaction of two scissors: mind and environment (Simon, 1956). To measure performance, the conception of ecological rationality is employed, i.e., what is the fit of the decision strategy with a given environment. A central departing point of the two traditions is the issue of learning. Heuristics and biases frequently point to situations where there have been no or little opportunities for learning, potentially resulting in a mismatch between heuristic and environment. In contrast, fast and frugal heuristics emphasize that strategies are evolved, i.e., there have been ample opportunities weeding out elements of a decision process or entire strategies that do not perform well.

Introduction

Firms are conventionally modeled as perfectly rational profit maximizers. This is also motivated by the claim that only those firms that act as-if they adhere to these principles survive. Yet, there is very little evidence about the actual decision process that firms employ for instance when setting prices. A recent interview study by Fabiani, Loupias, Martins, & Sabbatini (2007) with 11,000 managers in the Euro areas suggests that about two third of the firms employ pricing strategies that are not in line with perfectly rational profit maximization. Would this suggest that even in markets with considerable competitive pressure and ample opportunities for learning, behavior emerges that is inherently biased?

In paper three[3] we conduct interviews with 55 car dealers and collect pricing data of 748 dealers from an online used car market for two types of cars. From this data we deduce the principle pricing strategies that car dealers use. We find that 56% dealers selling 71% of used cars employ the 'step-price' heuristic: start with a high price and lower it in fixed time intervals until a car sells. Such a strategy ignores readily available information from the highly transparent online market where about 90% of used cars are listed (Dudenhöffer & Schadowski, 2011). Due to infrequent updating the strategy leads to price rigidity and price dispersion.

The step-price strategy stands in contrast to the conventional assumptions made in Friedman's (1953) as-if tradition in which a firm uses all information to compute its best response. Yet, it is not maladaptive. In fact it even outperforms a competitive strategy that updates prices frequently aiming to undercut competitors. The step-price strategy works well in the given environment due to an uncertain demand that dealers face which is fueled by consumers' uncertainty about the true quality of a car.

Summary

Friedman (1953) put forth that it is sufficient to predict outcomes and that this does not hinge on the underlying assumptions. The three studies presented here collectively suggest that understanding decision making requires to take into

[3] Artinger, F.. Step price heuristic: Frugal information use in the used car market.

account the interaction between mind and environment as originally envisioned by Simon (1956). Stable dispositions or preferences are likely to account only to a limited extend for behavior. Particularly the more complex an environment gets, the greater is the necessity to understand which elements of an environment shape behavior and which ones do not, thus contributing to the ecological rationality of a decision strategy. Decision processes that are more in line with the conventions of the as-if approach, for instance using all the available information, do not necessarily outperform simpler strategies. Analogously to the decision processes of expert chess players, in the face of a complex or uncertain environment, learning and competitive pressure can give rise to well adapted, robust decision strategies. These in turn can inform economic theories potentially yielding greater predictive power.

Bibliography

Berg, N., & Gigerenzer, G. (2010). As-If Behavioral Economics: Neoclassical Economics in Disguise? *History of Economic Ideas*, *18*(1), 133 – 166.

Camerer, C., & Loewenstein, G. F. (2004). Behavioral economics: Past, present, and future. In C. Camerer, G. F. Loewenstein, & M. Rabin (Eds.), *Advances in behavioral economics* (pp. 3–52). Princeton University Press.

De Groot, A. D. (1978). *Thought and choice in chess* (Vol. 4). Mouton De Gruyter.

Dudenhöffer, F., & Schadowski, A. (2011). Die Online-Gebrauchtwagen-Börsen in Europa. Universität Duisburg-Essen, CAR-Center Automotive Research.

Fabiani, S., Loupias, C. S., Martins, F., & Sabbatini, R. (2007). *Pricing Decisions in the Euro Area: How Firms Set Prices and Why*. Oxford University Press.

Fehr, E., Schmidt, K. M., Kolm, S.-C., & Ythier, J. M. (2006). The Economics of Fairness, Reciprocity and Altruism - Experimental Evidence and New Theories. *Foundations* (Vol. 1, pp. 615–691). Elsevier.

Friedman, M. (1953). *Essays in positive economics*. University of Chicago Press.

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science*, *15*(8), 534–539.

Iturbe-Ormaetxe, I., Ponti, G., Tomás, J., & Ubeda, L. (2011). Framing effects in public goods: Prospect Theory and experimental evidence. *Games and Economic Behavior*, *72*(2), 439–447.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291.

Luce, D. R. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, *24*(2), 178–191.

McCarter, M. W., Rockmann, K. W., & Northcraft, G. B. (2010). Is it even worth it? The effect of loss prospects in the outcome distribution of a public goods dilemma. *Organizational Behavior and Human Decision Processes*, *111*(1), 1–12.

Paal, T., & Bereczkei, T. (2007). Adult theory of mind, cooperation, Machiavellianism: The effect of mindreading on social relations. *Personality and Individual Differences*, *43*(3), 541–551.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138.

Simon, H. A., & Schaeffer, J. (1992). The Game of Chess. In R. J. Aumann & S. Hart (Eds.), *The Handbook of Game Theory with Economic Applications Vol. 1* (Vol. 1, pp. 1–17). Amsterdam: North Holland Press, Elsevier Science Publishers.

Singer, T., & Fehr, E. (2005). The Neuroeconomics of Mind Reading and Empathy. *American Economic Review*, *95*(2), 340–345.

Singer, T., & Lamm, C. (2009). The social neuroscience of empathy. *Annals of the New York Academy of Sciences*, *1156* (The Year in Cognitive Neuroscience 2009), 81–96.

Smith, V. L. (2008). *Rationality in Economics: Constructivist and Ecological Forms* (1st ed.). Cambridge University Press.

Takagishi, H., Kameshima, S., Schug, J., Koizumi, M., & Yamagishi, T. (2010). Theory of mind enhances preference for fairness. *Journal of Experimental Child Psychology*, *105*(1-2), 130–137.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124 –1131.

Putting yourself in others' shoes: Are fairness concerns shaped by individual differences in theory of mind and empathy?[1]

FLORIAN ARTINGER [a,*], FILIPPOS EXADAKTYLOS [b],

HANNES KOPPEL [c], LAURI SÄÄKSVUORI [c]

[a] Max Planck Institute for Human Development

[b] University of Granada

[c] Max Planck Institute of Economics

This study investigates how individual differences in the capacity for theory of mind (ToM) and empathy shape fairness concerns. We find no evidence that differences in empathy relate to fairness concerns in a neutral Dictator and Ultimatum Game. Using these games and accuracy of beliefs as a proxy for ToM, we show that fair individuals have overall more accurate beliefs about the action of others confirming previous results. This has invited speculations that higher ToM promotes fair behavior and therefore is an element underlying social preference theories. However, in situations where the accurate assessment of what others do is essential to further one's own material interest as in the Ultimatum Game, fair and selfish participants perform equally well in estimating the actions of others. This results in a steep positive correlation between accuracy of beliefs as a proxy for ToM and earnings for selfish participants. Our results qualify the role of ToM as a genuine source of fair behavior and question in how far individual differences in ToM and empathy are part of the psychological foundation of social preference theories.

Keywords: Theory of Mind; Empathy; Individual Difference; Personality; Altruism; Social Preferences.

JEL: C91, D03, D83

* *Corresponding author:* Florian Artinger, Max Planck Institute for Human Development, Center for Adaptive Behavior and Cognition, Lentzeallee 94, 14195 Berlin, Germany, Phone: +49 30 82406 699, artinger@mpib-berlin.mpg.de

---

[1] Contributions of the authors are as following: Idea: FA; Development of experimental design: FA, assisted by FE, HK, LS; Programming of experiments and data collection: HK, LS; Data analysis: HK, LS; Writing of manuscript: FA, assisted by FE, HK, LS.

Individual differences in fairness concerns

A large body of evidence indicates that a substantial share of people do not only care about their own material gains but are also motivated by fairness concerns. Yet, there are individual differences in how strongly people are motivated by such concerns. These insights, to a large extend from neutrally framed laboratory games, have had a decisive impact on the emergence of social preference theories (e.g., Fehr & Schmidt, 1999; Loewenstein, Thompson, & Bazerman, 1989). Subsequent research indicates that while theories like the inequity aversion model by Fehr & Schmidt (1999) capture some aggregate behavior across situations, they have little predictive power on the individual level (Blanco, Engelmann, Koch, & Normann, 2010). In social interaction, theory of mind (ToM) and empathy are regarded as central to guiding behavior (e.g., Binmore, 1990; Byrne & Whiten, 1988; Eisenberg & Miller, 1987; Kirman & Teschl, 2010; McCabe, Smith, & LePore, 2000; Singer & Fehr, 2005). In this paper, we investigate whether individual differences in ToM and empathy can account for the observed heterogeneity in fairness in neutrally framed laboratory games.

ToM and empathy relate to different areas of social decision making: (i.) social reasoning, how the other person is likely to act in a given situation; and (ii.) social emotions, emotional responses to, for instance, unfair or fair decisions (Singer, 2008). ToM is the capacity to understand the social reasoning and social emotions of others. Empathy is the capacity to share the social emotions of others.

To investigate individual differences in concerns for fairness, two laboratory games are particularly prominent: the Ultimatum Game (Güth, Schmittberger, & Schwarze, 1982) and Dictator Game (Kahneman, Knetsch, & Thaler, 1986).[2] In the Ultimatum Game (UG), a proposer states how much she wants to offer from an initial endowment. A responder can either accept or reject. If she rejects, both will get nothing. A proposer that seeks to earn the highest payoff must infer what the responder is likely to minimally accept. Due to its combination of strategic interaction and fairness concerns, the UG is particularly well positioned to evaluate the relationship between ToM and fairness. In the DG, the responder

[2] The prominence of these two games can be seen, for instance, with regards to the work of Fehr & Schmidt (1999) who use in their model on inequity aversion both games to estimate the distribution of fairness concerns in a population.

must accept any offer. That is why the DG is considered as a pure measure of individual concerns for fairness in the absence of strategic considerations (Forsythe, Horowitz, Savin, & Sefton, 1994). Given the nature of the decision environment, a number of researchers have speculated about the importance of individual differences in empathy for fairness as observed in neutrally framed laboratory games (e.g., Andreoni & Rao, 2010; Binmore, 1990; Singer & Fehr, 2005).

Capacities in social decision making

Theory of mind

Utilizing the capacity for ToM, the decision maker constructs the mental states of others' making inferences about beliefs, intentions, and emotions (for a review see Mason & Macrae, 2008). The concept of ToM has initially been proposed by primatologists who suggest that it emerged due to social challenges of living in larger coalitions (Byrne & Whiten, 1988; Premack & Woodruff, 1978). ToM is regarded to have two functions. First, it facilitates cooperation as shown for instance in non-human primates where the capacity for ToM can lead to acts of spontaneous helping (Warneken & Tomasello, 2006). Second, ToM enables deception of others for personal gains, a function often labeled as 'Machiavellian intelligence' (Byrne & Whiten, 1988).

Surprisingly, to date there exists only empirical evidence for humans of the first function, that individual differences in ToM are positively correlated with cooperativeness. Contradicting the second function, i.e., the relationship of ToM and Machiavellian disposition, there is evidence that people high on the Machiavellian scale score low on the capacity for ToM (Lyons, Caldwell, & Shultz, 2010; Paal & Bereczkei, 2007). Using games, individual differences in ToM have only been studied among children: in the UG, Autists between the age of six and ten, who are impaired in ToM, are more likely to accept low offers and to refuse fair proposals compared with a normally developed cohort (Sally & Hill, 2006). Likewise, pre-schoolers who have acquired the capacity for ToM make higher mean offers in the UG than those who have not yet developed this capacity (Takagishi, Kameshima, Schug, Koizumi, & Yamagishi, 2010). This has led to the proposition that differences in ToM are at the foundation of social preference

theories, whereby those with a higher ToM are also acting fairer (Takagishi et al. 2010). This conclusion stands in stark contrast to the second function and conventional assumption made in economics when modeling human behavior that strategic thinking informs the pursuit of one's material interest.[3]

Beliefs about the behavior of others' likely action are an inherent part of the capacity for ToM. In a seminal study on the role of beliefs in social dilemmas, Kelley & Stahelski (1970) show that cooperative individuals are more accurate in predicting the actions of others than selfish ones. They argue that this difference is functional in nature: for cooperative people to pursue their strategy successfully and reduce the likelihood of being suckers it is necessary to have more accurate beliefs than for defectors. This finding is in line with recent literature outlined above stating that higher ToM enhances the preference for fairness. It suggests that even absent the strategic necessity of a specific situation fair people can more accurately predict what others are likely to do. We refer to how closely the beliefs of a person match the action of another person as 'accuracy of beliefs'.

HYPOTHESIS 1: COMPARING FAIR AND SELFISH PARTICIPANTS AS MEASURED BY OFFERS IN THE DG, FAIR PARTICIPANTS HAVE MORE ACCURATE BELIEFS ABOUT OTHERS' DECISIONS THAN THE SELFISH ONES.

ToM and the accuracy with which one judges what others do plays an important role for those that aim to obtain the highest possible payoff for themselves in the UG: the better the person can estimate what others are likely to minimally accept, the higher is the earning that can be expected from the game. This implies that ToM does not only enhance a preference for fairness, as found in the recent literature, but that a higher ToM also enables higher earnings for more selfish people.

HYPOTHESIS 2: FOR SELFISH PARTICIPANTS, AS MEASURED BY THE OFFERS IN THE DG, THERE IS A POSITIVE CORRELATION BETWEEN ACCURACY OF BELIEFS AND EARNINGS IN THE UG.

[3] Strategic thinking is a necessary component for the maximization of personal gains, contrasting fair behavior.

Individual differences in ToM are usually measured using psychometric scales. A higher ToM score in these tasks should also facilitate better inferences about what others do in the context of games.

HYPOTHESIS 3: THE CAPACITY FOR TOM POSITIVELY CORRELATES WITH THE ACCURACY OF BELIEFS.

Empathy

Empathy is the capacity to share feelings with others. Individual differences in the capacity for empathy are shown to reflect differences in pro-social behavior in domains such as volunteering and donating (Davis et al. 1999). In a group of young adults, measures of pro-social dispositions have been found to be stable across five years and relate to ratings of empathy (Eisenberg et al. 2002). According to the empathy-altruism hypothesis, empathy is even regarded as the exclusive source of genuine altruism (Batson, Duncan, Ackerman, Buckley, & Birch, 1981). Empathic feelings are classically associated with helping someone in need (Eisenberg & Fabes, 1992; M. L. Hoffman, 2001). In the context of the Prisoner's Dilemma, inducing empathy via portraying the other person in need, results in participants choosing to cooperate more often than in a neutral control condition, even when subjects know that the other person will defect on them anyway (Batson & Moran, 1999).

Building on that argumentation, a number of researchers have speculated about the importance of individual differences in empathy for fairness as observed in neutrally framed laboratory games (e.g., Andreoni & Rao, 2010; Binmore, 1990; Singer & Fehr, 2005). Singer (p.264, 2008a) asserts that "one prediction that can easily be made is that people with a greater ability to empathize should display more other-regarding behavior". However, to date no research has investigated the question in how far differences in the disposition for empathy shape fairness concerns in neutrally framed laboratory games.

HYPOTHESIS 4: THE CAPACITY FOR EMPATHY POSITIVELY CORRELATES WITH OFFERS IN THE DG.

Method

Behavioral measures

In order to measure an individual's concern for fairness, participants made decisions in three roles[4]: as a proposer in the UG and DG and as a responder in the UG. In the role of the responder, participants indicated the minimum acceptance level below which they reject an offer. Participants were informed that they were randomly matched for each role with a different partner. At the end of the experiment, one of the three roles was picked randomly to calculate the payoffs.

In both games, the proposer was endowed with 90 currency units and could split these in intervals of 10. Equal splits are known to be the modal offers in the UG and DG. Hence, our design implied that participants were not able to offer an equal split removing the focal points of the games. It has been shown that after removing the equal split, fair offers become less frequent (Güth, Huck, & Müller, 2001) and that responders are less averse to unequal outcomes (Dana, Cain, & Dawes, 2006). This provides an interesting test bed for individual differences in ToM.

Participants were asked to indicate their beliefs about the likely action of a randomly chosen partner in each of the three decision tasks (DG offer, UG offer and UG minimum acceptance level).[5] Participants thereby stated the probability they considered most likely for a given action answering the question: "Please indicate the likelihood that a randomly determined person taking part in this experiment has chosen one of the ten possible divisions [minimum acceptance levels]". For instance, in the UG participants assigned a probability how likely a randomly matched partner would make any one of the ten offers from 0 to 90.

[4] When participants must indicate their actions in all possible roles this has been labeled the strategy vector method (Selten, 1967). Comparing this to a standard game where participants played only one role little difference was found (Oxoby & McLeish, 2004).

[5] Eliciting beliefs and actions of the same subjects in an incentive compatible manner, subjects might hedge across the different tasks, i.e., in one task they act more risky to increase their chance of higher earnings but compensate this on another task. Blanco, Engelmann, Koch, & Normann (2010) find that unless hedging opportunities are very prominent, results are not confounded. Since subjects did not know about the belief elicitation until they were actually asked to state their beliefs, prominence of hedging opportunities is not an issue here.

This resulted in a complete probability distribution of beliefs for a given action (for details see in the appendix).

Beliefs about the action of others were rewarded based on the Quadratic Scoring Rule (QSR) (Murphy & Winkler, 1970). The rule guarantees incentive compatibility and is widely used to elicit beliefs in games (for an extensive discussion on belief elicitation using the QSR see Artinger, Exadaktylos, Koppel, & Sääksvuori, 2010). To ensure full comprehension of the payoff mechanism, participants were carefully instructed on the procedure. Each participant had to go through three learning episodes with control questions that became increasingly more difficult. Overall, 96 percent of the participants were able to correctly answer the most demanding question. A higher QSR score, which is equivalent to a higher payoff, implies greater accuracy between the participant's belief and the actually observed action. The assessment of every participant is matched with each of the observed actions, resulting in 115 separate QSR. From this, we compute for every participant a mean score for the three tasks and an overall score that aggregates over the three separate tasks. The possible accuracy of belief scores range from 0 to 2.

The incentivization of beliefs circumvents a problem that Lyons et al. (2010) in their study on ToM and Machiavellianism point out: those participants high on the Machiavellian scale showed a lack of effort in the ToM task, indicating that these participants might not be well motivated by non-incentivized tests to reveal their ToM level. Using an incentivized proxy for the ability of ToM, we circumvent this limitation.

Psychometric measures

Ferguson, Heckman, & Corr (2011) link individual differences in preferences that have been traditionally subject in economics with individual differences in personality traits and dispositions. ToM and empathy have been traditionally assessed using psychometric measures for dispositions to investigate individual differences. We use two psychometric tests that both measure empathy and ToM. The two empathy measures are convergent whereas the two ToM measures complement each other: (i) "cold" ToM, tests for inferences with regards to the epistemic state of others (ii) "hot" ToM relates to others' emotions (Stone, 2000).

To measure cold ToM and empathy, we employed the Interpersonal Reactivity Index (IRI)[6] which has been extensively investigated and validated (Davis, 1980). The modified version used here has four dimensions each containing four statements. Two dimensions are used for the final analysis: cold ToM is measured using the 'perspective taking scale' where participants responded to statements like "I try to look at everybody's side of a disagreement before I make a decision". The capacity for empathy is measured with the 'empathic concern' scale. Participants had to respond to statements like "I often have tender, concerned feelings for people less fortunate than me". Individual differences in the empathic concern scale correlate positively with brain activity associated with empathy (Singer et al. 2004). Participants indicated on a 5-point scale in how much a statement "does not describe me well" to "describes me very well". Cronbach's α is 0.73 for the ToM measure and 0.80 for the empathy measure which corresponds to the values of Davis (1980).

To measure hot ToM and empathy, the Multifaceted Empathy Test (MET) is used which employs 40 realistic photographs of faces expressing positive or negative emotions as stimuli (Dziobek et al. 2007). Participants answered for each picture three types of questions reflecting three subscales. The subscale 'emotion recognition' measures hot ToM by assessing in how far participants can correctly infer the emotional state of others as depicted on the photographs. Participants answered the question "What does this person feel?" by selecting one out of four possible options where only one was correct. A similar test for hot ToM, relying on emotion recognition in faces, was used by Paal & Bereczkei (2007) and Lyons et al. (2010) who found a positive relationship between cooperativeness and the capacity for ToM. To measure empathy, the MET provides two subscales: the subscale 'direct empathy' asked participants to answer the question "How much do you feel with this person?". The scale 'indirect empathy' asked the question "How aroused are you by the picture?". Direct and indirect empathy are measured with a 9-point scale ranging from 'not at all' to 'a lot'. Cronbach's α are 0.95 for direct empathy, 0.96 for indirect empathy, and 0.70 for hot ToM which corresponds to the values found by Dziobek et al. (2007).

[6] The German translation of the IRI, the Saarbrücker Persönlichkeits-Fragebogen, was used in the experiment (Gigli & Paulus, 2009).

Procedure

A total of 120 students took part in one of four experimental sessions. They were recruited via the online recruitment system ORSEE (Greiner, 2004). The computerized experiment took place in the laboratory of the Max Planck Institute of Economics using z-Tree (Fischbacher, 2007). At the beginning of a session, each participant was seated in an isolated booth. All decisions were made anonymously. One session lasted about two hours. Earnings per participant ranged from 7 Euro to 26 Euro with a mean of 14 Euro. Four participants had to be excluded as they were not native German speakers, limiting their ability to fully respond to the verbal descriptions of emotional states used in one of the psychometric test.

Half of the participants played the DG, then the UG, followed by the psychometric tests; the other half of participants played first the psychometric tests and then the games. We do not find any order effects between the psychometric tests and games. We therefore pool the data. Subsequently, participants indicated their beliefs about the likely action of others in the games. As a fourth element, participants' risk attitude was measured using the Holt & Laury (2002) task in an incentive compatible manner where participants choose ten times between two different lotteries that gradually vary the combination of probabilities and outcomes. Questions on the demographic background of each participant ended the experiment.

Results

Behavioral data

Mean offers increase from 25 percent in the DG to 40 percent in the UG. Modal offers in both games are equally high at 44 percent. The mean of the minimum acceptance level in the UG is 26 percent, the mode is 33 percent.

To test our first hypothesis that fair participants have more accurate beliefs than selfish ones we use a mean split of DG offers in order to distinguish between fair and selfish participants. Alternative definitions of selfishness (e.g., median split or offering nothing defining the group of selfish participants) yield qualitatively very similar results to the here applied mean split. The accuracy of beliefs is, as

described above, measured by participant's ability to assess the behavior of the fellow participants. The scores for each of the three tasks and the respective test statistics are summarized in table 1. In support of our first hypothesis, comparing the beliefs of fair and selfish participants, we find that fair participants evince a higher accuracy of beliefs about offers made in the DG and UG as compared to selfish participants. Thus hypothesis 1 is supported for beliefs about offers made in the DG and UG. However, there is no difference in accuracy of beliefs between fair and selfish participants for the minimum acceptance level in the UG. A possible explanation for this result is that being able to estimate the minimum acceptance level well is of particular relevance for selfish participants who seek to incur a high payoff.

TABLE 1. MEAN ACCURACY OF BELIEFS (STANDARD DEVIATION). ACROSS ALL THREE TASKS INCLUDING THE TOTAL SCORE FOR SELFISH (N = 58) AND FAIR (N = 58) PARTICIPANTS. MANN-WHITNEY U TEST AND TWO-TAILED ASYMPTOTIC P-VALUES ARE SHOWN.

	Fair Mean (SD)	**Selfish** Mean (SD)	**MWU**	**p-value**
DG offer	1.07 (0.12)	0.95 (0.20)	1050.50	0.00
UG offer	1.23 (0.11)	1.14 (0.21)	1259.00	0.00
UG min. acceptance	1.05 (0.15)	1.05 (0.21)	1484.50	0.28

Testing hypothesis 2 that accuracy of beliefs and monetary earnings of selfish participants should positively correlate, we estimate the expected monetary earnings for each participant from the UG (hypothesis 2). First, we calculate how often a participant's offer in the UG is accepted. This is done by comparing participant's UG offer with the minimally acceptance level of each of the 115 responders. We label this a participant's expected acceptance rate. Multiplying the expected acceptance rate with the respective participant's offer, we form an indicator of the expected monetary earnings for each participant.

Fair participants' offers are accepted with 94 percent. Selfish participants' mean acceptance rate is 80 percent due to the lower offers they make. This leads to similar earnings of selfish and fair participants: the mean payoff from the UG offer for selfish participants is 46.68; fair participants earn 43.70 monetary units in the mean (Mann-Whitney U (58, 58) = 1350.50, p = .08).

As figure 1 shows, there is a high positive correlation (r = 0.75, p = .02) for selfish participants between accuracy of beliefs about the minimum acceptance level and expected earnings in the UG. There is no significant correlation between the accuracy of beliefs and expected earnings in the UG for fair participants (r = -0.14, p = .21). This finding supports hypothesis 2 and indicates that selfish participants effectively utilize ToM for their own monetary benefit; the higher their accuracy of beliefs the higher their earnings and the higher the inequality between the proposer and responder. Particularly for the selfish participants, being able to accurately estimate what others are likely to minimally accept reduces the risk of rejection and increases expected earnings.

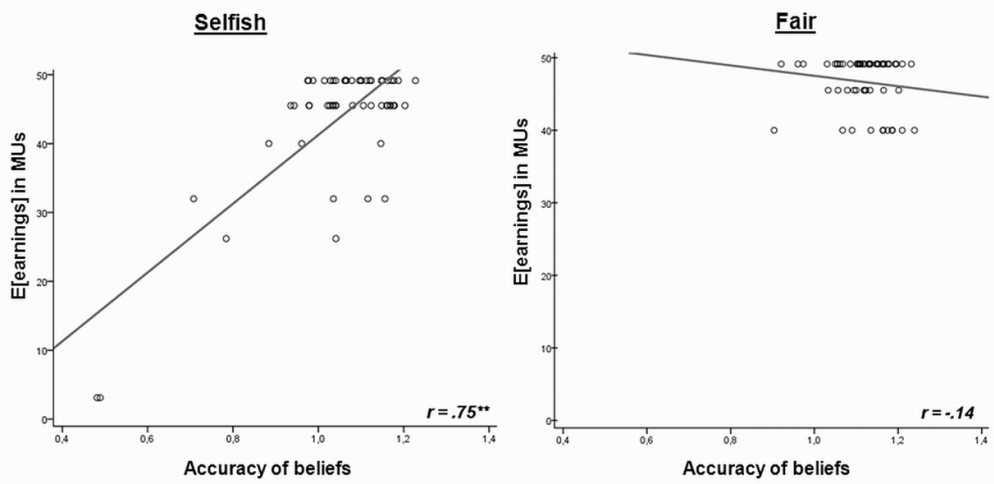

FIGURE 1. EXPECTED EARNINGS IN THE UG GIVEN THE ACCURACY OF BELIEFS ABOUT UG MIN. ACCEPTANCE FOR SELFISH AND FAIR PARTICIPANTS.

Behavioral and psychometric data

An important question both from a behavioral and methodological perspective is if the capacity for ToM, as measured by IRI and MET, positively correlates with the accuracy of beliefs (hypothesis 3). As shown in table 2 below, ToM and the accuracy of beliefs are not significantly correlated regardless of the psychometric instrument. This is a surprising result particularly for cold ToM.

TABLE 2. PEARSON'S CORRELATION COEFFICIENTS (P-VALUES) FOR MET AND IRI SCALES WITH ACCURACY OF BELIEFS AND OFFERS, N=116.

	IRI - cold ToM	MET - hot ToM
Accuracy of beliefs UG min acceptance	-0.08 (0.42)	-0.04 (0.67)
Accuracy of beliefs DG offer	0.06 (0.51)	-0.03 (0.77)
Accuracy of beliefs UG offer	-0.05 (0.57)	-0.05 (0.63)
Accuracy of beliefs mean TOTAL	-0.03 (0.72)	-0.06 (0.55)
DG offer	0.20 (0.04)	-0.10 (0.29)
UG offer	0.08 (0.41)	0.03 (0.73)

Furthermore, we find a positive correlation between cold ToM and offers in the DG, but no relationship to offers in the UG. This suggests that as one moves from a unilateral, non-strategic, setting (DG) to a bilateral, strategic, interaction (UG) cold ToM as measured by the IRI does not relate to behavior anymore.

Why do the accuracy of beliefs and cold ToM not correlate? One possible explanation is that the two tests for ToM (accuracy of beliefs and IRI – cold ToM) are addressing different environments such as a strategic vs. a non-strategic context. Asking people about beliefs what others are likely to do could focus behavior on strategic interaction. However, the accuracy of beliefs in the DG is not related to cold ToM. These results indicate that the IRI (and the MET) might not be well suited to relate to behavior in an anonymous, strategic context.

Do offers in the DG and UG correlate with measures for empathy (hypothesis 4)? We find no significant positive correlation. Table 3 shows correlations coefficients between offers in the games and the empathy scores measured by the psychometric tests. Correlating empathy scores with various alternative behavioral indexes of altruism (for instance, UG offer minus UG belief minimum acceptance; belief of DG offer by fellow participants minus DG offer) does not uncover any statistically significant relation.

TABLE 3. PEARSON'S CORRELATION COEFFICIENTS (P-VALUES) FOR MET AND IRI SCALES FOR EMPATHY WITH OFFERS, N=116.

	IRI - Empathy	MET - Direct Empathy	MET - Indirect Empathy
DG offer	0.12 (0.19)	-0.08 (0.42)	-0.06 (0.53)
UG offer	0.10 (0.28)	-0.08 (0.40)	-0.06 (0.56)

Discussion

In this paper we ask how individual differences in ToM and empathy influence fairness concerns. We find that fair participants have more accurate beliefs about offers made in the DG and UG, confirming previous results. However, when it comes to judging what others would minimally accept in the UG this difference disappears. We suggest that this result can be explained by accurately estimating the minimum acceptance level being particularly relevant for selfish participants who seek to incur a high payoff: accurate beliefs about the minimum acceptance level are essential to choosing the lowest possible offer that would be accepted thus allowing maximizing payoffs for the proposer in the UG. Indeed we find a steep positive correlation between accuracy of beliefs as a proxy for ToM and earnings for selfish participants.

Our result is in line with Lyons et al. (2010) who suggest that the lower scores of ToM found in Machiavellians as compared to non-Machiavellians (Ali & Chamorro-Premuzic, 2010; Lyons et al. 2010; Paal & Bereczkei, 2007) are due to a lack of incentivization of selfish participants. Different than the psychometric measures used in these studies, our approach offers monetary incentives to the participants showing that selfish participants indeed utilize ToM effectively for their own means.

In fact, the psychometric measures for empathy and ToM show little relation to the behavioral data in our study. This puts doubt on the link between (social) preferences on the one hand and dispositions on the other hand (e.g., Ferguson et al. 2011; Takagishi et al. 2010). Particularly individual differences in empathy have been speculated to be related to fairness concerns as observed in a neutrally framed laboratory game (e.g., Andreoni & Rao, 2010; Binmore, 1990; Singer, 2008; Singer & Fehr, 2005). Using two convergent psychometric measures, one relying on pictures as stimuli (MET) the other on verbal descriptions (IRI), both do not indicate that there is any relationship between fairness concerns and empathic disposition in the given context. This suggests that the empathy-altruism hypothesis (Batson et al. 1981), whereby empathy is regarded as the exclusive source of genuine altruism, might not hold under more general conditions. Rather, it finds support in a suitably framed environment as for instance in Batson &

Moran (1999) where a person is framed in need. This suggests that social emotions in a neutrally framed laboratory game do not feature prominently to elicit any empathic response.

For the ToM measures of the psychometric tests, we find a weak positive correlation between DG offers and cold ToM, which tests for inferences with regards to the epistemic state of others, but not hot ToM which relates to inferences about others' emotions (Stone, 2000). This further supports the conjecture from the measures on empathy that social emotions do not feature prominently here. The weak positive correlation between DG offers and cold ToM might suggest that cold ToM is employed to match social expectations by more fair participants. However, at the same time, the proxy for cold ToM, the accuracy of beliefs, does not correlate with the psychometric measure. In studies with children it has been observed that ToM positively correlates with offers in the UG (Sally & Hill, 2006; Takagishi et al. 2010). In contrast to our study, these games were not played anonymously but an experimenter was paired with each child and was present at all times. Such a design might invite a demand effect where participants want to appear fair as suggested by a number of recent studies on the DG and UG that indicate that greater anonymity or the possibilities to obscure the role of the proposer decreases offers (for a review see Cooper & Kagel, in prep.).

Our results put the role of ToM as a fundamental psychological foundation of social preference theories into perspective and invites future research addressing in which environments certain measures of ToM are expected to perform well, potentially focusing on differences between unilateral, non-strategic, settings (DG) and bilateral, strategic interactions (UG).

Bibliography

Ali, F., & Chamorro-Premuzic, T. (2010). Investigating Theory of Mind deficits in nonclinical psychopathy and Machiavellianism. *Personality and Individual Differences*, *49*(3), 169–174.

Andreoni, J., & Rao, J. M. (2010). The Power of Asking: How Communication Affects Selfishness, Empathy, and Altruism. *NBER Working Papers*.

Artinger, F., Exadaktylos, F., Koppel, H., & Sääksvuori, L. (2010). *Applying Quadratic Scoring Rule transparently in multiple choice settings: A note*. Friedrich-Schiller-University Jena, Max-Planck-Institute of Economics.

Batson, C. D., Duncan, B., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathetic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology*, *40*, 290–302.

Batson, C. D., & Moran, T. (1999). Empathy-induced altruism in a prisoner's dilemma. *European Journal of Social Psychology*, *29*(7), 909–924.

Binmore, K. G. (1990). *Essays on the Foundations of Game Theory*. Blackwell Pub.

Blanco, M., Engelmann, D., Koch, A. K., & Normann, H. T. (2010). Belief elicitation in experiments: is there a hedging problem? *Experimental Economics*, 1–27.

Byrne, R. W., & Whiten, A. (1988). *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford University Press, USA.

Cooper, D. J., & Kagel, J. H. (n.d.). Other-regarding preferences. In J. H. Kagel & A. E. Roth (Eds.), *The Handbook of Experimental Economics* (Vol. 2).

Dana, J., Cain, D. M., & Dawes, R. M. (2006). What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, *100*(2), 193–201.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, *10*(4), 85.

Davis, M. H., Mitchell, K. V., Hall, J. A., Lothert, J., Snapp, T., & Meyer, M. (1999). Empathy, Expectations, and Situational Preferences: Personality Influences on the Decision to Participate in Volunteer Helping Behaviors. *Journal of Personality*, *67*(3), 469–503.

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H., Wolf, O., & Convit, A. (2007). Dissociation of Cognitive and Emotional Empathy in Adults with Asperger Syndrome Using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders*, *38*(3), 464–473.

Eisenberg, N., & Fabes, R. A. (1992). Emotion, regulation, and the development of social competence. In M. Clark (Ed.), *Review of personality and social psychology* (pp. 119 – 151). Thousand Oaks: Sage Publications, Inc.

Eisenberg, N., Guthrie, I. K., Cumberland, A., Murphy, B. C., Shepard, S. A., Zhou, Q., & Carlo, G. (2002). Prosocial development in early adulthood: A longitudinal study. *Journal of Personality and Social Psychology*, *82*(6), 993–1006.

Eisenberg, N., & Miller, P. A. (1987). The Relation of Empathy to Prosocial and Related Behaviors. *Psychological Bulletin*, *101*(1), 91–119.

Fehr, E., & Schmidt, K. M. (1999). A Theory Of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics*, *114*(3), 817–868.

Ferguson, E., Heckman, J. J., & Corr, P. (2011). Personality and economics: Overview and proposed framework. *Personality and Individual Differences*, *51*(3), 201–209.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in Simple Bargaining Experiments. *Games and Economic Behavior*, *6*(3), 347–369.

Gigli, M., & Paulus, C. (2009). *Der Saarbrücker Persönlichkeitsfragebogen SFP (IRI) zur Messung von Empathie: Psychometrische Evaluation der deutschen Version des Interpersonal Reactivity Index*. Working paper.

Greiner, B. (2004). *The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics*. University of Cologne, Department of Economics.

Güth, W., Huck, S., & Müller, W. (2001). The Relevance of Equal Splits in Ultimatum Games. *Games and Economic Behavior*, *37*(1), 161–169.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An Experimental Analysis of Ultimatum Bargaining. *Journal of Economic Behavior and Organization*, *3*(4), 367–388.

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., et al. (2010). Markets, Religion, Community Size, and the Evolution of Fairness and Punishment. *Science*, *327*(5972), 1480–1484.

Hoffman, E., McCabe, K. A., & Smith, V. L. (1996a). On expectations and the monetary stakes in ultimatum games. *International Journal of Game Theory*, *25*(3), 289–301.

Hoffman, E., McCabe, K., & Smith, V. L. (1996b). Social Distance and Other-Regarding Behavior in Dictator Games. *The American Economic Review*, *86*(3), 653–660.

Hoffman, M. L. (2001). *Empathy and Moral Development: Implications for Caring and Justice*. Cambridge University Press.

Holt, C. A., & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review*, American Economic Review, *92*(5), 1644–1655.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the Assumptions of Economics. *The Journal of Business*, *59*(4), 285–300.

Kelley, H. H., & Stahelski, A. J. (1970). Social interaction basis of cooperators' and competitors' beliefs about others. *Journal of Personality and Social Psychology*, *16*(1), 66–91.

Kirman, A., & Teschl, M. (2010). Selfish or selfless? The role of empathy in economics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1538), 303 –317.

Loewenstein, G., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social psychology*, *57*(3), 426–441.

Lyons, M., Caldwell, T., & Shultz, S. (2010). Mind-reading and manipulation — Is Machiavellianism related to theory of mind? *Journal of Evolutionary Psychology*, *8*(3), 261–274.

Mason, M. F., & Macrae, C. N. (2008). Perspective-Taking from a Social Neuroscience Standpoint. *Group Processes & Intergroup Relations*, *11*(2), 215 –232.

McCabe, K. A., Smith, V. L., & LePore, M. (2000). Intentionality detection and "mindreading": Why does game form matter? *Proceedings of the National Academy of Sciences of the United States of America*, *97*(8), 4404 –4409.

Murphy, A. H., & Winkler, R. L. (1970). Scoring rules in probability assessment and evaluation. *Acta Psychologica*, *34*, 273–286.

Oxoby, R. J., & McLeish, K. N. (2004). Sequential decision and strategy vector methods in ultimatum bargaining: evidence on the strength of other-regarding behavior. *Economics Letters*, *84*(3), 399–405.

Paal, T., & Bereczkei, T. (2007). Adult theory of mind, cooperation, Machiavellianism: The effect of mindreading on social relations. *Personality and Individual Differences*, *43*(3), 541–551.

Premack, D., & Woodruff, G. (1978). Chimpanzee problem-solving: a test for comprehension. *Science*, *202*(4367), 532–535.

Sally, D., & Hill, E. (2006). The development of interpersonal strategy: Autism, theory-of-mind, cooperation and fairness. *Journal of Economic Psychology*, *27*(1), 73–97.

Selten, R. (1967). Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments. *Beiträge zur experimentellen Wirtschaftsforschung*, *1*, 136–168.

Singer, T. (2008). Understanding others: Brain mechanisms of theory of mind and empathy. In P. W. Glimcher, C. F. Camerer, R. A. Poldrack, & E. Fehr (Eds.), *Neuroeconomics: Decision Making and the Brain* (pp. 251 – 268). London: Academic Press.

Singer, T., & Fehr, E. (2005). The Neuroeconomics of Mind Reading and Empathy. *American Economic Review*, *95*(2), 340–345.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science*, *303*(5661), 1157–1162.

Stone, V. E. (2000). The role of the frontal lobes and the amygdala in theory of mind. In S. Baron-Cohen, H. Tager-Flusberg, & D. Cohen (Eds.), *Understanding Other Minds: Perspectives from Autism* (2nd ed., pp. 253–273). Oxford: Oxford University Press.

Takagishi, H., Kameshima, S., Schug, J., Koizumi, M., & Yamagishi, T. (2010). Theory of mind enhances preference for fairness. *Journal of Experimental Child Psychology*, *105*(1-2), 130–137.

Warneken, F., & Tomasello, M. (2006). Altruistic Helping in Human Infants and Young Chimpanzees. *Science*, *311*(5765), 1301–1303.

Appendix

Instructions

Thank you for coming! You are now about to take part in an experiment on decision making. By reading carefully the following instructions and taking part in the experiment you can earn money depending both on your own decisions and on the decisions of others.

These instructions and the decisions to be made are solely for your private information. During the experiment, you are not allowed to communicate in the laboratory or with someone outside the laboratory. Please switch off your mobile phone. Any violation of these rules will lead to your exclusion from the experiment and no payments. If you have any questions regarding the rules or the course of this experiment, please raise your hand. An experimenter will assist you privately.

This experiment consists of one computerized questionnaire and three separate sections with varying decision tasks. Answering carefully all items in the questionnaire will earn you 4 Euro. In each of the three separate sections, one randomly chosen decision determines your earnings from the section. Your overall income from the experiment will be based on the sum of earnings from the three separate sections and the questionnaire. Neither during nor after the experiment will you or any other participant be informed about the identity of a person with whom you are interacting. Your earnings will be paid privately in cash at the end of the experiment.

During the experiment, all decisions are made in Experimental Currency Units (ECUs). Your total income will be calculated in ECUs and at the end of the experiment converted to Euros at the following rate: 1 ECU = 0.15 Euro

The experiment begins with the questionnaire. You have one decision to be taken per computer screen. Please bear in mind that after the introductory stage of two computer screens you have up to 15 seconds to make your decision on each screen. The remaining time is displayed on your screen in the upper right hand corner.

First section

The first section consists of two decision tasks in which your earnings depend both on your own decisions and one randomly chosen participant. There are two types of individuals: Type A and type B. You will make a decision in both roles. To calculate your earnings from this section, only one decision will be randomly chosen. The random decision is determined by the computer at the end of the experiment.

First decision task

There are two types of individuals: Type A and type B. Person A decides how to divide a pie of 90 ECUs between him/herself and a person B. Person B is passive in this situation. The division is possible only in intervals of 10 currency units. In other words, person A can allocate 0, 10, 20, 30, 40, 50, 60, 70, 80, or 90 ECUs to person B.

Example: Person A allocates 30 ECUs to person B, person A earns (90 - 30 =) 60 ECUs.

Second decision task

Person A decides how to divide a pie of 90 ECUs between him/herself and person B. Person B may now either accept or decline the proposed division. Should person B accept the division, both persons earn ECUs in compliance with the proposed division. Should the person B decline the offered allocation, both persons earn nothing. To determine the final allocation from the second decision task, person B indicates the minimum amount of ECUs that he/she is willing to accept. Person A's decision to divide the pie and B's minimum that he/she is willing to accept are only possible in intervals of 10 currency units. You are now asked to make your decision in both roles: as person A and B. Your payoff relevant decision will be randomly determined by the computer at the end of the experiment.

Example: Person A decides to offer 30 ECUs to person B and thereby keeps 60 ECUs for him/herself. Person B indicates the minimum amount of ECUs he/she is willing to accept. Should the amount be smaller or equal to 30 ECUs, receives

person B 30 ECUs and person A 60 ECUs. Should the acceptable amount be greater than 30 ECUs, do both persons receive 0 ECUs from the decision task.

Second section

The second section consists of three decision tasks that are described below. In this section, your earnings depend both on your own decisions and on the decisions of others. At the end of the experiment, the computer will randomly determine one of the three tasks that will be used to determine your earnings from the second section.

In the first stage there were three situations in which 90 ECUs were at stake:

1. Person A allocates ECUs between him/herself and person B, B is passive;
2. Person A allocates ECUs between him/herself and person B, B is active;
3. Person B indicates the smallest amount that he/she is willing to accept.

All decisions were to be made in intervals of 10 ECUs.

In the following three decision tasks, your earnings will be determined by the accuracy of your probability assessment. Your task is to indicate the likelihood that a randomly determined person has chosen one of the ten possible divisions / minimum acceptance of ECUs. Please note that the sum of your probability assessments needs to equal 100 per cent. Your earnings will be calculated based on the following figure. A more detailed explanation will follow.

Stated probability (in percent)	Choice of partner correctly predicted (in ECU*)	Costs for giving probabilities to not chosen action (in ECU*)
100%	20,00 ECU	10,00 ECU
95%	19,98 ECU	9,03 ECU
90%	19,90 ECU	8,10 ECU
85%	19,78 ECU	7,23 ECU
80%	19,60 ECU	6,40 ECU
75%	19,38 ECU	5,63 ECU
70%	19,10 ECU	4,90 ECU
65%	18,78 ECU	4,23 ECU
60%	18,40 ECU	3,60 ECU
55%	17,98 ECU	3,03 ECU
50%	17,50 ECU	2,50 ECU
45%	16,98 ECU	2,03 ECU
40%	16,40 ECU	1,60 ECU
35%	15,78 ECU	1,23 ECU
30%	15,10 ECU	0,90 ECU
25%	14,38 ECU	0,63 ECU
20%	13,60 ECU	0,40 ECU
15%	12,78 ECU	0,23 ECU
10%	11,90 ECU	0,10 ECU
5%	10,98 ECU	0,03 ECU
0%	10,00 ECU	0,00 ECU

Note: * ECU stands for Experimental Currency Unit

The payoff consequences of your choices will be explained through the following example: Assume a situation in which person A decides how to allocate a pie of 90 ECUs between him/herself and a person B. Person B is passive. The first column in the table contains the probability that you want to assign for a certain possible division. Should you for instance assess that all 10 possible divisions (from 0 ECUs to 90 ECUs) are equally likely to occur, your decision is to set 10 per cent probability to all possible events.

The second column in the table indicates your earnings from a correct prediction given your probability assessment. In the example, all 10 possible divisions received a probability estimate of 10 per cent. You have inevitably made a correct prediction, which earns you 11.90 ECUs.

You have to bear the costs from incorrect probability assessments (third column). In this example, you have set 10 per cent probability also for all the events that did not occur. These incorrect predictions are all associated with a deduction 0.10 ECUs as can be read from the third column in the table.

That is, your total earnings from the task are 11.90 ECU - 0.10 ECU - 0.10 ECU - 0.10 ECU - 0.10 ECU - 0.10 ECU - 0.10 ECU - 0.10 ECU - 0.10 ECU - 0.10 ECU - 0.10 ECU = 11.90 ECU -- 9 * 0.10 ECU = 11.00 ECU.

Further example: Assume that you have made following probability assessments: 10% for 0 ECUs, 20% for 10 ECUs, 40% for 20 ECUs and 15% each for 30 and 40 ECUs.

The randomly chosen person A decides to allocate 10 ECUs to person B. Your probability assessment for that event was 20%. Your earnings from the decision task will be calculated as following: 13.60 ECUs (20% for a correct prediction) - 0.10 (10% for an incorrect prediction) - 1.60 ECUs (40% for an incorrect prediction) - 2 * 0.23 ECUs (two times 15% for an incorrect prediction) = 11.44 ECUs.

Please pay attention to the fact that under the given payoff scheme the worst possible monetary outcome happens when you set 100 per cent probability for an event that does not occur. Your earnings in such case would be 0 ECUs. On the contrary, should you set 100 per cent probability for an event that occurs, your earnings would be the highest possible with 20 ECUs. Note that you are not bound to make your probability assessments in intervals of 5 percent. This limitation was used only for an illustration. That is, you can for instance set a probability of 97% for a certain event. You will receive a complete payoff table once we begin the experiment.

Training Period

Next we will ask you questions regarding the decision situation. The questions will help you to understand the calculation of your payoff and ensure that you have understood the instructions. Once you have found the correct answer, please insert the value in the given box.

When you have answered both questions correctly, we are ready to begin the experiment. Please answer the following questions with the help of the above given table.

1. You are expecting with a probability of 55% that your partner will give you 30 ECUs. Your partner decides to actually give you 30 ECUs. How many ECUs will you receive for this correct prediction?

2. You are expecting with a probability of 50% that your partner will decide to give you 30 ECUs, with a probability of 30% 20 ECUs, and with a probability of 20% that he will decide to give you 10 ECUs. Your partner decides to actually give you 40 ECUs. How high are your costs (in ECUs) for this incorrect assessment?

3. You are expecting with a probability of 50% that your partner will decide to give you 30 ECUs, with a probability of 30% that he will give you 40 ECU, with a probability of 10% that he will give you 20 ECUs, with a probability of 5% that he will give you 10 ECUs, and with a probability of 5% that he will give you 50 ECUs. Your partner decides to actually give you 30 ECUs. How high is your payoff (in ECUs)?

Cooperation in a risky environment: decisions from experience in a stochastic social dilemma[1]

FLORIAN ARTINGER [a,*], NADINE FLEISCHHUT [a],
M. VITTORIA LEVATI [b,c], JEFFREY R. STEVENS [a,d]

[a] Max Planck Institute for Human Development

[b] Max Planck Institute of Economics

[c] University of Verona

[d] University of Nebraska-Lincoln

Cooperation is often subject to environmental risk. We investigate how cooperation is shaped by the way information about risk is presented (from description or from experience) and by differences in the degree of environmental risk. Drawing on research from risky choice, we compare choices in stochastic social dilemmas to those in lotteries with an equivalent degree of environmental risk. In stochastic games with the same expected outcomes, cooperation rates vary with different degrees of risk, mimicking decision patterns observed in lotteries. Risk presentation, however, only affects choices in lotteries, not in stochastic games. Process data suggest that people respond less to probabilities in the stochastic social dilemmas than in the lotteries. These findings highlight how elements of uncertainty in the environment shape cooperation.

Keywords: Decisions from Experience; Social Dilemma; Risky Choice; Public Good.

JEL: C72, C73, C92, D81

* *Corresponding author:* Florian Artinger, Max Planck Institute for Human Development, Center for Adaptive Behavior and Cognition, Lentzeallee 94, 14195 Berlin, Germany, Phone: +49 30 82406 699, artinger@mpib-berlin.mpg.de

[1] Contributions of the authors are as following: Idea: FA; Development of experimental design: NF, FA, assisted by JS, VL; Programming of experiments and data collection: NF; Data analysis: FA; Writing of manuscript: FA, NF, assisted by VL, JS.

Cooperation in a risky environment

When people face an opportunity to cooperate, such as when opening a business together or pursuing a joint research project, the outcomes of these enterprises are often risky. This risk can have two sources: social and environmental. Joint enterprises often constitute a social dilemma, where it is in the collective interest of the group to cooperate, yet individually rational to free ride. Despite these incentives, there is overwhelming evidence that people frequently engage in cooperation (Ostrom, 1990; Ridley, 1996). Yet, even if people cooperate, outcomes are often subject to a risky environment. For instance, even if all business partners would cooperate, a new start-up may fail due to external events such as the demand breaking down after a macroeconomic shock. Laboratory experiments show that when social dilemmas are embedded in a risky environment which influences outcomes, cooperation declines sharply (for a review see E. Van Dijk, Wit, Wilke, & Budescu, 2004). What has not been addressed is how the way in which information about such environmental risks is presented and how the degree of environmental risk affect cooperation.

A recent flurry of studies document a pronounced difference in behavior depending on how information is presented: whether people sample the distribution of outcomes (*decisions from experience*), or decide based on a description of outcomes and probabilities (*decision from description*) (Hertwig, Barron, Weber, & Erev, 2004; for a review see Hertwig & Erev, 2009; Rakow & Newell, 2010). In conventional lotteries with described probabilities, people choose as-if they overestimate the impact of small probabilities as reflected in Prospect Theory (PT) (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). In contrast, people decide as-if they underestimate the impact of small probabilities if they acquire information on the riskiness of a situation sequentially by sampling (Hertwig et al. 2004). The difference in choice patterns between decisions from description and experience has been labeled the description-experience gap (DE gap). The difference between decisions from description and experience can in fact be traced to Knight (1921). He suggested that there are (a) a priori probabilities where events result from precisely known stochastic mechanism for which a probability can be assigned; (b) statistical probabilities where the mechanism is not known and probabilities are assessed in an empirical

manner; (c) estimates where no probability, neither from a mechanism nor from empirical assessment can be deduced, commonly referred to as uncertainty (Hau, Pleskac, & Hertwig, 2010).

People's risky choices vary with *the degree of risk*, i.e. with different combinations of outcomes and probabilities while the expected value is constant. Analogously, one can introduce a stochastic component into a social dilemma retaining defection as the dominant strategy. Differences in outcomes have been shown to affect the rate of cooperation in a one-shot, deterministic, Prisoner's Dilemma (Guyer & Rapoport, 1972). In this paper, we do not only vary outcomes but also the probability with which these occur addressing whether and how differences in the degree of risk affect the rate of cooperation in a one-shot social dilemma.

Our understanding of cooperation in stochastic environments is currently limited to situations in which environmental risk is described by outcomes and probabilities (e.g., Bereby-Meyer & Roth, 2006; Gangadharan & Nemes, 2009; Gong, Baron, & Kunreuther, 2009; Kunreuther, Silvasi, Bradlow, & Small, 2009; Levati, Morone, & Fiore, 2009). We argue that real-world risky choices often involve experiencing the outcome and probabilities of choices rather than receiving their summary statistics. Therefore, examining how risk presentation influences people's decisions is critical to understand how a stochastic environment shapes cooperation.

Prospect Theory, which has been originally proposed to describe risky choice, has recently been applied to a stochastic social dilemma: choices in a stochastic public good's game (PG) show that the degree of cooperation depends on whether there is a probability of a loss or not (Iturbe-Ormaetxe, Ponti, Tomás, & Ubeda, 2011; McCarter, Rockmann, & Northcraft, 2010). The authors of both studies interpret this as loss aversion, a core proposition of PT. Underlying PT is a weight-and-add framework where people are modeled as-if they integrate and non-linearly transform all outcomes and probabilities. This is evident in the non-linear weighting of probabilities in PT where small probabilities are overweighted and large probabilities are underweighted compared to their objective value. We evaluate this proposition in the context of a stochastic social dilemma by using a novel approach to PT that ranks Prospect Theory Values. This allows to compare

whether differences in the degree of risk are treated analogously in lotteries and matching social dilemmas.

Like other types of choices, cooperation is a function of the match between decision processes and the decision-making environment, or what has been labeled ecological rationality (Gigerenzer, Todd, & the ABC Research Group, 1999; Simon, 1956; Smith, 2008). While social uncertainty has been studied extensively, the impact of environmental risk is critical for many real-world social interactions. Only recently researchers started to address this. For instance, Bereby-Meyer & Roth (2006) investigate the role of learning and find that cooperation unravels quicker over time in a deterministic than in a stochastic Prisoner's Dilemma. Gong, Baron, & Kunreuther (2009) compare cooperation between groups and individuals, finding that groups cooperate more in a stochastic Prisoner's Dilemma compared to the deterministic setting. Kunreuther, Silvasi, Bradlow, & Small (2009) show that people in a stochastic Prisoner's Dilemma do not only respond to what their counterpart did but also to whether their partner had suffered a loss in the last round. None of the studies, however, addresses how differences in risky environments and the way information about risk is presented affect cooperation.

Design

We employ a 2 x 2 between-subjects design in which we manipulate the risk presentation (description vs. experience) and the choice situation (social dilemma vs. lottery). In the description condition, subjects receive information about how environmental risk influences outcomes in terms of described probabilities, whereas in the experience condition participants sample to infer probabilities. The environmental risk is identical between lotteries and games. To investigate how differences in the degree of risk affect behavior in one-shot social dilemmas, we vary probabilities and outcomes within-subjects while keeping the expected outcomes constant.

Environmental risk in social dilemmas and lotteries

We use a stochastic 2-person public goods game (PG) with binary choices. This is equivalent to a Prisoner's Dilemma except that participants are not shown a

payoff matrix but informed that for each choice they receive an endowment *e* which they can contribute to a joint project with a randomly matched partner or keep it for themselves. Contributions are multiplied by a value, the marginal social return (msr)[2], and shared equally between both players. Denoting *i*'s contribution by c_i, where $c_i \in \{0, e\}$ and $i = 1,2$, *i*'s payoff is given by:

$$\pi_i = e - c_i + \frac{msr}{2}(c_1 + c_2). \qquad (1)$$

We impose msr $\in \{1,2\}$. An msr that is larger than 1 makes it socially optimal to contribute; an msr smaller than 2 renders free-riding the dominant strategy, thus creating a social dilemma.

The msr can take on two possible values, representing either a good or a bad event, with a certain probability. In case the bad event occurs, contributions multiply by an msr < 1, decreasing the value of the contributions. When the good event occurs, contributions multiply by an msr > 1, increasing the value of the contributions. The environmental risk only affects what is invested. We chose the two msr-values and corresponding probabilities such that the expected msr, E[msr], across good and bad event always yields a social dilemma with 1 < E[msr] < 2.

Table 1 illustrates the eight decision situations that participants were confronted with. Situations 1 to 4 contain each one rare (p < .25) bad event, analogous to the DE gap studies with lotteries (e.g., Hertwig et al. 2004). Situations 5 and 6 contain only common (p > .25) events to test whether the DE gap extends beyond rare events as found by Ludvig and Spetch (2011). We use two different expected msr, 1.2 and 1.4, to check the robustness of the results. Situations 1 – 6 are designed to extend the findings from the DE gap studies in risky choice to social dilemmas. Keeping the expected msr constant allows us to test whether different combinations of probabilities and outcomes affect choices in the PG in the same

[2] Public goods games conventionally use the marginal per capita return (mpcr), i.e., what each participant earns from one unit of investment to the public good. The msr is equal to mpcr*n where n is the number of interacting subjects. Using the msr helps in our setting where (as will be clarified later) the msr might be lower than 1. In this case, the msr has the advantage that values below 1 are more easily understood as earnings that are potentially smaller than the initial endowment and values above 1 as earnings that are potentially larger than the initial endowment. Using the msr thus facilitates the comprehension by the subjects of the impact of the msr on earnings.

way that they affect choices in lotteries. Decision situation 7 and 8 explore extreme conditions of a social dilemma and provide a further control whether participants understand the incentives. In situation 7, the E[msr] equals 1.1, which makes it less attractive to cooperate compared to situations 1 – 6. In contrast to the other situations, the rare event is the good state of the world. In situation 8 the expected msr of 2.1 does not generate a social dilemma and makes it individually and socially optimal to cooperate.

In most studies on the DE gap, the risky option has an expected value that is only marginally higher than the sure option.[3] To avoid floor effects in the social dilemma, we use a relatively large expected msr. This should provide strong incentives to cooperate in the PG.

In lotteries, participants also receive an endowment e and have to decide whether to invest in a risky option. If they invest, this gets multiplied by one of two possible values with a certain probability, which are identical to the msr and their probabilities in the games. While the realized payoffs in the lotteries only depend on the realized state of the world, the realized payoffs in the games also depend on the action of another player. The lotteries strip the strategic component away, but retain the stochastic component that stems from the environment.

Consider decision situation 1 as a lottery: if a participant invests her endowment, she can earn either 1.30*e with 92 % probability or 0*e with 8 % probability; alternatively, if she does not invest, she earns e for sure. If decision situation 1 is a game and the good (bad) state is realized, each receives 1.30*e (0*e) if both pair members invest and e if both do not invest.

All parameters were selected to facilitate easy computation by the participants. Across decision situations, the expected msr-values were only equal up to the first decimal. The initial endowment e is 10 in both lotteries and games for each decision situation. We randomized the order of decision situations in games as

[3] See for instance the set of lotteries used by Hertwig, Barron, Weber, and Erev (2004) which have also been used widely by others. Here, the risky option is only about 1.07 times higher than the safe option. However, Hau, Pleskac, and Hertwig (2010) employ in lottery 1 an EV for the risky option that is 1.26 times larger than the safe option and find a DE gap if participants draw a larger number of samples.

well as in lotteries, and participants receive no feedback about the realized msr or decision of the other group member after each decision.

TABLE 1. DECISION SITUATIONS

Nr.	Lottery			Prisoner's Dilemma											
	Risky option	Expected msr		Expected msr				Good state				Bad state			
				C		NC		C		NC		C		NC	
Rare events															
1	1.30, 92% 0, 8%	1.2	**C**	1.2	1.2	0.6	1.6	1.3	1.3	0.7	1.7	0.0	0.0	0.0	1.0
			NC	1.6	0.6	1.0	1.0	1.7	0.7	1.0	1.0	1.0	0.0	1.0	1.0
2	1.45, 80% 0, 20%	1.2	**C**	1.2	1.2	0.6	1.6	1.5	1.5	0.7	1.7	0.0	0.0	0.0	1.0
			NC	1.6	0.6	1.0	1.0	1.7	0.7	1.0	1.0	1.0	0.0	1.0	1.0
3	1.55, 92% 0, 8%	1.4	**C**	1.4	1.4	0.7	1.7	1.6	1.6	0.8	1.8	0.0	0.0	0.0	1.0
			NC	1.7	0.7	1.0	1.0	1.8	0.8	1.0	1.0	1.0	0.0	1.0	1.0
4	1.80, 80% 0, 20%	1.4	**C**	1.4	1.4	0.7	1.7	1.8	1.8	0.9	1.9	0.0	0.0	0.0	1.0
			NI	1.7	0.7	1.0	1.0	1.9	0.9	1.0	1.0	1.0	0.0	1.0	1.0
Two common events															
5	1.80, 64% 0.20, 36%	1.2	**C**	1.2	1.2	0.6	1.6	1.8	1.8	0.9	1.9	0.2	0.2	0.1	1.1
			NC	1.6	0.6	1.0	1.0	1.9	0.9	1.0	1.0	1.1	0.1	1.0	1.0
6	1.95, 56% 0.70, 44%	1.4	**C**	1.4	1.4	0.7	1.7	2.0	2.0	1.0	2.0	0.7	0.7	0.4	1.4
			NC	1.7	0.7	1.0	1.0	2.0	1.0	1.0	1.0	1.4	0.4	1.0	1.0
Extreme expected msr															
7	0.75, 88% 3.50, 12%	1.1	**C**	1.1	1.1	0.5	1.5	3.5	3.5	1.8	2.8	0.8	0.8	0.4	1.4
			NC	1.5	0.5	1.0	1.0	2.8	1.8	1.0	1.0	1.4	0.4	1.0	1.0
8	2.20, 96% 0.30, 4%	2.1	**C**	2.1	2.1	1.1	2.1	2.2	2.2	1.1	2.1	0.3	0.3	0.2	1.2
			NC	2.1	1.1	1.0	1.0	2.1	1.1	1.0	1.0	1.2	0.2	1.0	1.0

Note. In the columns that provide a description of the risky option, each msr is followed by its probability of occurrence and the expected msr. In the game matrices, C stands for contributing, NC for not contributing. In each cell, the left (right) number denotes the payoff of the row (column) player.

Decisions by description vs. decision by experience

The main aim of the study is to investigate how risk presentation – either as description or experience – affects decisions to cooperate. In the description condition, participants received information about environmental risk as a summary statement about probabilities and associated msr-values before they made their decision. In the experience condition, participants sampled the distribution of msr-values by drawing 25 cards from a deck. We used a matched

sampling design based on Ungemach, Chater, & Stewart (2009), where people are forced to view a representative sample of the underlying distributions of outcomes. Each card contains a number corresponding to one of the two possible msr. For example, in situation 1 the deck has 2 cards with the msr of 0 and 23 cards with an msr of 1.30. The sequence of cards is randomized for each participant, yet the two possible msr-values and their frequencies match exactly the objective probabilities given in the description condition. Thus, a sampling error cannot cause any differences observed between the two conditions. In addition, we collect time stamps in the experience conditions which allow us to evaluate how long participants view a certain card and whether this influences their decision.

Risk sensitivity in social dilemmas and lotteries

The results of lotteries and stochastic social dilemmas are also evaluated with respect to predictions of PT. PT retains the original weight-and-add framework of Expected Value Theory to model human decision making. It thereby uses a separate value and weighting function to transform the expected value into Prospect Theory Values (PTVs). A different tradition of modeling human behavior is represented by process theories which aim to capture the decision process that people actually employ.[4] The priority heuristic (PH) of Brandstatter, Gigerenzer, & Hertwig (2006) proposes a lexicographic strategy to evaluate lotteries where decision makers are modeled as sequentially evaluating elements of the decision environment instead of integrating these all at the same time. It can predict empirical data better than modifications of EUT such as Cumulative Prospect Theory (Tversky & Kahneman, 1992).[5]

Ideally, competing theories from both traditions would have been used to make predictions about the decision situations in games and lotteries. However, only a novel way of using PT can tackle the issues at hand. Lotteries and stochastic PGs

[4] Pioneering work in risky choice has been done by Rubinstein (1988) and Leland (1994).

[5] The priority heuristic contains the following steps: 1. Priority rule: Go through the attributes of lotteries in the order of: minimum gain, probability of minimum gain, maximum gain. 2. Stopping rule: stop information search if the minimum gains differ by one tenth or more of the maximum gain (across lotteries), otherwise stop search if probabilities of minimum gain differ by .1 or more. 3. Decision rule: Choose the lottery that is more attractive in the attribute which stopped search.

contain the same stochastic element. However, investment rates in the PG are expected to be lower than in lotteries due to a second source of uncertainty that stems from the interaction with the other person. Conventionally, PT and PH have been used to make predictions about majority choices. This does not allow evaluating whether behavior in lotteries and games responds in the same way to changes in the degree of risk. Yet, PTVs also produce a ranking which of eight decision situations should receive more investment. Such a ranking does not rest on majority predictions but highlights the relative rate of investment and can be applied to both lotteries and stochastic PGs in the description condition. Table 2 lists the PTVs for the eight decision situations in this experiment based on the parameters used by Tversky and Kahneman (1992). For decision situations that have the same expected msr (1,2 and 5 on the one hand and 3,4, and 6 on the other hand), the ranking yields predictions with regards to how people would change their behavior facing differences in the degree of risk. This ranking is denoted by romantic numbers for E[msr] = 1.2 from I – III and by letters for E[msr] = 1.4 from A-C.

TABLE 2. PREDICTIONS OF PROSPECT THEORY.

	Risky Option	Expected msr	Prospect Theory	
			PTVs	Ranking
Rare event				
1	1.30, 92% 0, 8%	1.2	0.93	II
2	1.45, 80% 0, 20%	1.2	0.84	III
3	1.55, 92% 0, 8%	1.4	1.09	B
4	1.80, 80% 0, 20%	1.4	1.02	C
Two common events				
5	1.80, 64% 0.20, 36%	1.2	0.96	I
6	1.95, 56% 0.70, 44%	1.4	1.21	A
Extreme expected msr				
7	0.75, 88% 3.50, 12%	1.1	1.23	
8	2.20, 96% 0.30, 4%	2.1	1.70	

Note. In the columns that provide a description of the risky option, each outcome is followed by its probability of occurrence and the expected value. The outcome of the sure option is always 1. The ranking is done for decision situations that have the same expected msr.

Further tasks

To check the accuracy of risk estimates in the experience conditions, after the last round we asked participants how often they have seen the two sampled msr-values. In the description condition, participants translate the probability statement of the last round into a frequency statement to control whether participants accurately understood the risk. In the social dilemma conditions, participants also face two deterministic PGs in randomized order after the eight stochastic situations: one with an msr of 1.2 (which matches the expected msr in situations 1, 2, and 5), the other with an msr of 1.4 (which matches the expected msr in situations 3, 4, and 6). By this means we can investigate how cooperation varies if the stochastic component is removed.

At the end of the experimental session, participants indicated which of six reasons best explains their decision to invest / not invest into the stochastic PGs. Responses were solicited via a questionnaire adapted from McCarter, Rockmann, & Northcraft, (2010). Participants could select one of the following six reasons

why they invested / did not invest (see appendix A2 for the complete questionnaire): the probability to get a low payoff was (not) sufficiently high, the two potential msr-values were (not) sufficiently high, conditional cooperation, social uncertainty, greed/opportunism, moral values, or none of these. A section on demographics concluded the experiment. All tasks are completed anonymously employing a perfect stranger design. At the end of the experiment, one decision situation is randomly chosen to determine final payouts.

Participants and procedure

128 students were recruited using the ORSEE (Greiner, 2004) software and were randomly assigned to one of four sessions: either to stochastic two-person PGs or to lotteries in either a description or an experience condition. In the social dilemma conditions, participants had to pass control questions to ensure that they understood the impact of environmental risk and of the other person's choice on their payoffs. Participants earned, on average, 14€. The sessions with lotteries lasted 60 minutes, those with games 75 minutes. The experiment was conducted at the laboratory of the Max Planck Institute of Economics in Jena using z-Tree (Fischbacher, 2007). No participant took part in more than one session, and participants were generally relatively unfamiliar with laboratory experiments.

Hypotheses

The goal of the study is to investigate how cooperation is affected by risk presentation and by the degree of risk, i.e. different combinations of outcomes and probabilities. PT is used as a benchmark. Based on table 2's PTVs, two predictions follow in decisions from description for both lotteries and games with the same expected msr:

> *Hypothesis 1a: Situations containing an outcome that is realized with 8 percent (1 and 3) receive more investment than situations containing an outcome with 20 percent (2 and 4).*
> *Hypothesis 1b: in decision situations with a common event (5 and 6) the rate of investment is higher than in decision situations with a rare event (1, 2 and 3, 4).*

Using lotteries, studies have demonstrated pronounced behavioral differences depending on whether information on risk is experienced or described. For experienced risk, people choose the risky option more often compared to

decisions from description. This has been modeled as if people underestimate the impact of small probabilities. Extending this choice pattern to social dilemmas leads to the following hypothesis: underestimating the impact of the less likely events in situations 1 to 6 (table 1), investment rates should be higher in experience than in description. This also holds for situation 8. This pattern should reverse for situation 7, where the less common event is the good outcome.

Hypothesis 2: In lotteries and the stochastic PGs, the risky option is more frequently picked in the experience condition than in the description condition. This pattern reverses for decision situation 7 where the good event is less likely.

Results

Risk sensitivity in social dilemmas and lotteries

How do differences in the stochastic environment affect cooperation? For the results of hypothesis 1a and 1b, we focus on data from the description condition for decision situations 1 to 6, where the same expected outcomes are implemented. Figure 1 illustrates the behavior in lotteries and stochastic PGs for the two expected msr of 1.2 and 1.4. For comparison, it includes the rate of investment in the deterministic PG. The x-axis displays the probability of the (bad) less likely payoff of a decision situation and the y-axis the percentage of participants who choose to invest.

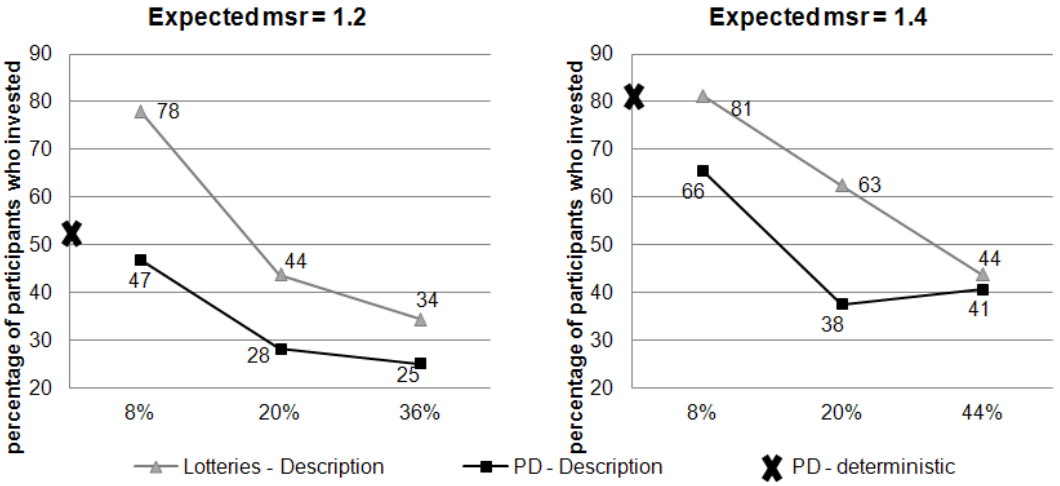

Figure 1. Investment rates in lotteries and PGs in description; level of investment in deterministic PG.

When comparing investment rates in the left and right panel of figure 1, cooperation rates are higher for larger expected msr. The deterministic PGs yield a similar pattern: the rate of cooperation is 53% with a sure msr = 1.2 and, 81% with a sure msr = 1.4 ($\chi^2(1) = 5.74$, $p = .02$). In the stochastic PGs, the mean rate of cooperation is 33% when E[msr] = 1.2 and 48% when the E[msr] = 1.4 ($\chi^2(1) = 4.23$, $p = .04$). Thus, differences in expected msr affect behavior even though the social dilemma is maintained and the dominant strategy for a person is not to cooperate. This replicates the finding of Guyer & Rapoport (1972) and extends it to a stochastic setting. But, besides being sensitive to different expected returns, do people react to differences in risk for constant expected msr?

To address this question, we pool the decision situations that contain the same probabilities across the expected msr-values (e.g., situations 1 and 3 and situations 2 and 4) to obtain more reliable results.[6] The mean cooperation rate is 1.7 times higher in situations where the bad event occurs with 8% than in situations where the bad event is occurs with 20% ($\chi^2(1) = 7.12$, $p = .01$). Thus, changes in the stochastic environment have a large impact on cooperation. Note that the difference in cooperation between deterministic PG and stochastic PG with an 8% chance of a bad event becomes quite small. In fact, this difference is not significant (pooling across the expected msr: $\chi^2(1) = 1.62$, $p = .20$).

To investigate hypotheses 1a and 1b, one can also rely on the pooled data across different expected msr (e.g., situations 1,3; situations 2,4; situations 5,6) because the rankings of PTVs are identical for situation with an expected msr of 1.2 and 1.4. The rate of investment in situations with a probability of 8% compared to 20% sharply drops both for stochastic PGs and lotteries (for stochastic PGs, the rate of investment declines in the pooled data from 56% to 33% ($\chi^2(1) = 7.17$, $p = .01$); for lotteries, the rate of investment declines in the pooled data from 80% to 53% ($\chi^2(1) = 10.12$, $p < .001$). Paralleling each other, stochastic PGs and lotteries thus are in line with prediction 1a based on Prospect Theory: situations with 8% receive more investment than situations with 20%.

[6] Earlier studies (Harless & Camerer, 1994; Hey & Orme, 1994; for a summary see Wakker, 2010) report a considerable noise in lotteries. Employing a repeated within-subject design, they find that participants are inconsistent in about one out of four choices.

For prediction 1b, Figure 1 also suggests a decline in cooperation between situations with a probability of 20% and those with two common events. Statistically, however, there is no difference between these two situations, neither for the stochastic PGs (the rate of investment is constant at 33%, $\chi^2(1) = .00$, $p = 1.00$), nor for lotteries (the rate of investment declines in the pooled data from 53% to 39%, ($\chi^2(1) = 2.55$, $p = .11$). However, hypothesis 2b based on Prospect Theory – that the rate of investment is highest with a common event – is neither met in stochastic PGs nor in lotteries.

In summary, we find that differences in the degree of environmental risk influence choice in the PGs for decisions from description. Differences in the stochastic environments result in similar behavior in stochastic PGs and lotteries. Though the data confirm the predictions of Prospect Theory for hypothesis 1a, they do not support hypothesis 1b for either PGs or lotteries.[7]

Decisions from description and from experience

We now address how risk presentation affects behavior in games and whether it does so in the same way as in lotteries (hypothesis 2). Following, we explore the reasoning processes in stochastic PGs in experience and description as compared to lotteries.

Is there a DE gap in lotteries and games?

Hypothesis 2 is directional and states that except for situation 7, participants should invest more in the experience condition. To test this hypothesis, we subtract the percentage of people contributing in the experience condition from those in the description condition, except for situation 7 where we do the opposite. Figure 2 illustrates the DE gap pooling across all eight decision situation for lotteries and stochastic PGs. The box plot lists the minimum, first quartile, median, third quartile, and maximum. The results show a positive gap for lotteries ($\chi^2(1) = 8.24$, $p = .003$), with a mean difference between experience and description of 12% (SD = 10%).

[7] For comparison, the constant-relative-risk-aversion utility of the decision situation can be calculated for instance by using the parameterization in Goeree, Holt, & Palfrey (2003). However, the predictions are also not met. The calculations for instance suggest that investment in decision situation (DS) 6 > DS 3 which is not met, neither in games nor lotteries.

Table 3 lists the investment rate in experience and description separately for all 8 decisions situations in lotteries and stochastic PGs. With regard to lotteries, the predicted difference between the description and experience condition is observed for all lotteries except for lottery 8 where the expected outcome is twice as high as the sure option. Averaging across lotteries containing a rare bad event (lotteries 1-4) shows a DE gap of 13% (Table 3). The same DE gap (13%) occurs with lotteries containing more common events (5 and 6, Table 3). The results replicate Ludvig and Spetch (2011), who find the DE gap also for situations with common events. Overall, supporting previous findings, responses to decisions from description and experience differed in lotteries in the predicted direction. Thus, the parameters we chose for environmental risk replicate the DE gap found in the risky choice literature.

Given that the parameters replicate the DE gap in lotteries and given that behavior in games is similarly sensitive to differences in risk than in lotteries as shown by the previous result, we expected the way risk is presented to also influence levels of cooperation. The behavior in the stochastic PGs, however, does in this respect not mimic the behavior in lotteries: the DE gap completely disappears in games (χ^2 (1) = .38, p = .30). The mean difference between experience and description in the stochastic PG is -3% (SD = 6).

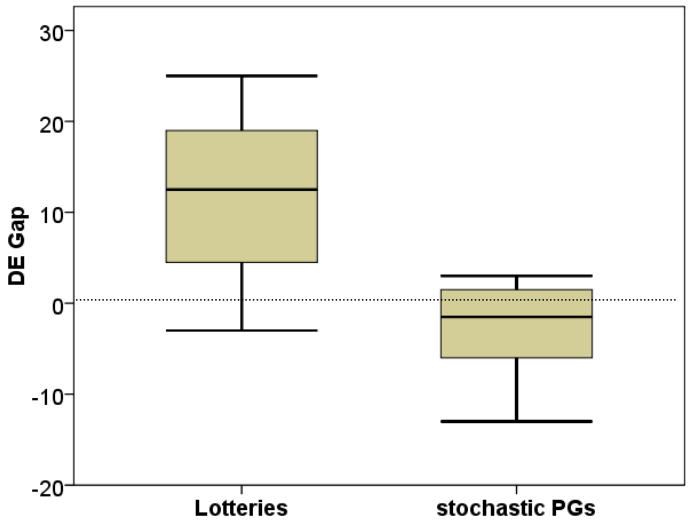

FIGURE 2. BOX PLOTS OF THE DE GAP POOLED ACROSS EIGHT SITUATIONS IN LOTTERIES AND GAMES.

The lotteries stand in stark contrast to the results in the stochastic PG. In games, 6 out of 8 decision situations show no or only minimal DE gaps (Table 3). Experience and description conditions do not differ for any of the decision situations. In fact, situation 7 which is closest in spirit to the situations used by Hertwig et al. (2004) and Ungemach et al. (2009) shows a strong DE gap in lotteries, but the gap disappears completely in the games.

TABLE 3. PERCENTAGE OF CHOICES IN INVESTMENT IN DESCRIPTION AND EXPERIENCE

	Risky Option	E[msr]	Lotteries			Stochastic PD		
			Experience	Description	Difference	Experience	Description	Difference
Rare event								
1	1.30, 92% 0, 8%	1.2	81	78	+3 ($\chi^2(1) = .10$, $p = .76$)	44	47	-3 ($\chi^2(1) = .06$, $p = .80$)
2	1.45, 80% 0, 20%	1.2	69	44	+25 ($\chi^2(1) = 4.06$, $p = .04$)	28	28	0 ($\chi^2(1) = .00$, $p = 1.00$)
3	1.55, 92% 0, 8%	1.4	88	81	+6 ($\chi^2(1) = .47$, $p = .49$)	56	66	-9 ($\chi^2(1) = .59$, $p = .44$)
4	1.80, 80% 0, 20%	1.4	78	63	+16 ($\chi^2(1) = 1.87$, $p = .17$)	38	38	0 ($\chi^2(1) = .00$, $p = 1.00$)
Mean 1 - 4			*79*	*66*	+13 ($\chi^2(1) = 5.03$, $p = .03$)	*41*	*45*	-3 ($\chi^2(1) = .26$, $p = .61$)
Two common events								
5	1.80, 64% 0.20, 36%	1.2	44	34	+9 ($\chi^2(1) = .59$, $p = .44$)	28	25	3 ($\chi^2(1) = .08$, $p = .77$)
6	1.95, 56% 0.70, 44%	1.4	59	44	+16 ($\chi^2(1) = 1.56$, $p = .21$)	28	41	-13 ($\chi^2(1) = 1.11$, $p = .29$)
Mean 5 & 6			*52*	*39*	+13 ($\chi^2(1) = 2.02$, $p = .16$)	*28*	*33*	-5 ($\chi^2(1) = .33$, $p = .57$)
Extreme expected msr								
7	0.75, 88% 3.50, 12%	1.1	16	38	-22 ($\chi^2(1) = 3.92$, $p = .05$)	16	19	-3 ($\chi^2(1) = .11$, $p = .74$)
8	2.20, 96% 0.30, 4%	2.1	97	100	-3 (*$p = .50$, Fisher's exact test*)	88	91	-3 ($\chi^2(1) = .16$, $p = .69$)

Note. In situation 8, due to a cell having only a count of 1, Fisher's exact test is applied, per condition n = 32.

Why is there a DE gap in lotteries but not in games?

Contrary to expectations, we observe a DE gap in lotteries but not games. In the following we investigate to what extend the decision process differs between lotteries and games giving a fist indication why the DE gap disappears.

One possible explanation underlying this pattern is that participants spend different amounts of time sampling lotteries and games which might indicate differences in attention given to the two conditions. In lotteries, participants spend more time viewing the rare event (M = .91 seconds, SD = .99) compared to the common event (M = .67 seconds, SD = .65, t(6,400) = 10.01, p < .000). Similarly, for the games, participants view the rare event (M = .51 seconds, SD = .51) longer than the common event (M = .43 seconds, SD = .33, t(6,400) = 6.38, p < .000). In lotteries, however, participants spent more time sampling than in games for both rare events (t(2,432) = 12.45, p < .000) and common events (t(10,368) = 24.02, p<.000). These differences in sampling times provide evidence that participants may pay less attention to the actually observed probabilities in the games as compared to lotteries with identical environmental risks.

FIGURE 3. MEAN SECONDS SPEND PER DRAW IN LOTTERIES AND GAMES FOR RARE EVENTS (R) AND COMMON EVENTS (C).

To control for the accuracy of risk perception, participants in the experience conditions stated the frequency of the two outcomes in the last situation after they had decided. The actual distribution of outcomes participants saw in their last round correlates with the stated frequencies for lotteries (r_S = .72, p < .000) yet to a lesser extent for stochastic PGs (r_S = .43, p < .01). These two correlations are not significantly different from each other at a conventional level (z = -1.73, p = .08). This suggests that in both conditions participants are calibrated to the actual

probabilities and did not underestimate but if at all rather overestimate the probability of rare events (see figure A1 in the appendix).

Some researchers suggest that the larger influence of recent events in decisions from experience may drive the DE gap. Hertwig et al. (2004) and Rakow, Demes, & Newell (2008) find a recency effect in decisions from experience but Ungemach et al. (2010) and Hau, Pleskac, Kiefer, & Hertwig (2008) do not. To test for a recency effect, we divide the 25 samples of the outcomes participants draw in our study into two sets: from 1 to 12 (initial) and from 13 to 25 (latter). Then we compute the expected msr from the initial samples, $E[msr]_{1-12}$, and from the latter samples, $E[msr]_{13-25}$. Finally, we compare the number of risky choices made when $E[msr]_{13-25} > E[msr]_{1-12}$ to the number of risky choices made when $E[msr]_{13-25} < E[msr]_{1-12}$ (see table A1 in the appendix). We find that recency plays a role in lotteries ($\chi^2(1) = 3.77$, $p = .04$) but not in games ($\chi^2(1) = .30$, $p = .34$). This also suggests that the actually observed probabilities may play a less important role in games than in lotteries.

Finally, for the stochastic PG in description and experience, figure 4 displays the reasons participants indicated as most important for their decision to cooperate and not to cooperate. This gives two statements per participants (see table A2 in appendix). Aggregating across both statements, when being confronted with the PG in description the probability in the stochastic games is the most important reason for 59 percent of participants, however, in the experience condition it is the most important reason for only 39 percent of the participants. In the experience condition, participants also emphasize the value of the msr they can obtain (20% in experience, and 3% in description) and conditional cooperation, i.e. their expectation whether the other will cooperate (20% in experience and 11% in description). This indicates that the importance of probabilities for decisions is further reduced in the stochastic PG in experience.

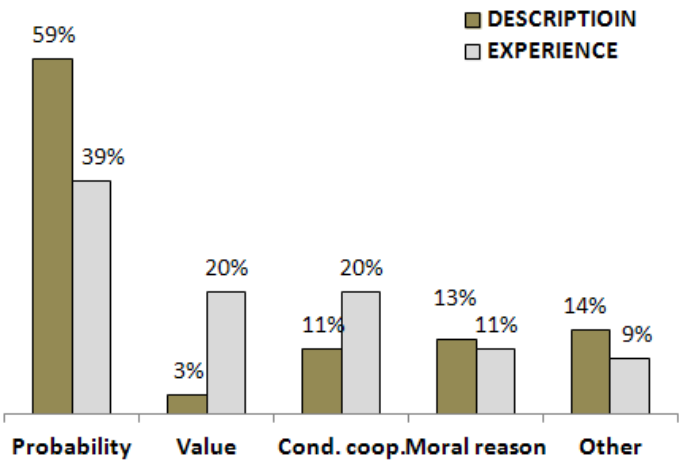

FIGURE 4. MAIN REASON STATED FOR INVESTING / NOT INVESTING IN STOCHASTIC PGS.

In summary, in the experience condition participants sample more quickly in the stochastic PG than in lotteries, suggesting that they are paying less attention to the observed probabilities. In line with this, subjects risk perception was less accurate in games than in lotteries, and recency — a potential cause of the DE gap — does not play a role in games, whereas we do find a recency effect in lotteries. Similarly, the questionnaire also highlights that probabilities are less important in the PG in experience than the size of the values and beliefs about others' behavior. This provides converging evidence that the probabilities of the risky option lose importance in the games; the DE gap is washed out.

Discussion

Cooperation in social dilemmas has been observed widely. We examine how critical aspects of the stochastic environment shape cooperation. First, the degree of environmental risk influences cooperation. Investments decisions in the stochastic PGs thereby mimic those observed in lotteries. Keeping expected outcomes constant, people invest more in a decision situation containing an 8% chance of a bad event occurring over a decision situation with a 20% chance of the bad event. Second, the values and probabilities chosen to implement environmental risk replicate the DE gap in lotteries. That is, people choose the risky option more often when experiencing the risky outcomes compared to when receiving summary descriptions. Our key finding is that risk presentation matters

in lotteries but not in games: no DE gap exists for the social dilemmas. Process data and subjects self-reported reasons for cooperation suggests that the disappearance of the DE gap in games may result from a decision process that emphasizes the size of the outcomes and expectations about others' behavior over outcome probabilities.

From a Prospect Theory perspective, choices in decisions from experience have been described as inversing the non-linear weighting of probabilities found in decisions from description, which leads to the gap in choices between both conditions (Hertwig et al. 2004; Ludvig & Spetch, 2011; Ungemach et al. 2009). In our study, however, PT could not even account for the data in the description condition for either lotteries or games. A potential reason may be that the stochastic environment chosen invites violations of stochastic dominance which PT cannot account for (Birnbaum, Coffey, Mellers, & Weiss, 1992). In particular, Wakker (2010) points out that 'zero' outcomes are liable to result in behavior not in line with PT. This shows the limits of introducing PT to a stochastic social dilemma (Iturbe-Ormaetxe et al. 2011; McCarter et al. 2010)

The results suggest that people do not integrate all the available information in such a way that adding a stochastic element to strategic interaction preserves the properties of decisions under risk. Instead certain elements become more important such as outcomes and strategic interaction whereas other elements loose importance, such as outcome probabilities. In order to understand more complex environments and thus move colser to the real world these results suggest that one needs to gain an understanding which elements of a decision environment dominate under what conditions. This also suggests that often lexicographic strategies might be a good approximation of human decision making (e.g., Luce, 1956). Models that focus on actual decision processes and their match with the environment may thus provide promising alternative to PT or other weight-and-add models in the tradition of Expected Value Theory (Gigerenzer et al. 1999; Simon, 1956; Smith, 2003). In fact, the basic message of our study – that outcomes feature more strongly than probabilities in some environments but not in others – has been pointed out elsewhere in the risky choice literature (e.g., Sunstein, 2003). The priority heuristic outlines a sequential decision process which considers outcomes in the first and probabilities only as a second step if no

decisions has been made (Brandstatter et al. 2006). Particularly with the use of process data, future research might be able to distil the actual decision process that people employ in social dilemmas that are also object to environmental uncertainty.

Bibliography

Bereby-Meyer, Y., & Roth, A. E. (2006). The Speed of Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation. *American Economic Review*, *96*(4), 1029-1042.

Birnbaum, M. H., Coffey, G., Mellers, B. A., & Weiss, R. (1992). Utility Measurement: Configural-Weight Theory and the Judge's Point of View. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 331–346.

Brandstatter, E., Gigerenzer, G., & Hertwig, R. (2006). The Priority Heuristic: Making Choices without Trade-Offs. *Psychological Review*, *113*(2), 409–432.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.

Ford, J. K., Schmitt, N., Schechtman, S. L., Hults, B. M., & Doherty, M. L. (1989). Process tracing methods: Contributions, problems, and neglected research questions. *Organizational Behavior and Human Decision Processes*, *43*(1), 75–117.

Gangadharan, L., & Nemes, V. (2009). Experimental analysis of risk and uncertainty in provisioning private and public goods. *Economic Inquiry*, *47*(1), 146–164.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, *62*, 451–482.

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press.

Goeree, J. K., Holt, C. A., & Palfrey, T. R. (2003). Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior*, *45*(1), 97–113.

Gong, M., Baron, J., & Kunreuther, H. (2009). Group cooperation under uncertainty. *Journal of Risk and Uncertainty*, *39*(3), 251–270.

Greiner, B. (2004). *The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics*. University of Cologne, Department of Economics.

Guyer, M. J., & Rapoport, A. (1972). 2× 2 Games Played Once. *The Journal of Conflict Resolution*, *16*(3), 409–431.

Harless, D. W., & Camerer, C. F. (1994). The Predictive Utility of Generalized Expected Utility Theories. *Econometrica*, *62*(6), 1251–1289.

Hau, R., Pleskac, T. J., & Hertwig, R. (2010). Decisions from experience and statistical probabilities: Why they trigger different choices than a priori probabilities. *Journal of Behavioral Decision Making*, *23*(1), 48–68.

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description-experience gap in risky choice: the role of sample size and experienced probabilities. *Journal of Behavioral Decision Making*, *21*(5), 493–518.

Hauser, J. (2011). A marketing science perspective on recognition-based heuristics (and the fast-and-frugal paradigm). *Judgment and Decision Making*, Judgment and Decision Making, *6*(5), 396–408.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from Experience and the Effect of Rare Events in Risky Choice. *Psychological Science*, *15*(8), 534–539.

Hertwig, R., & Erev, I. (2009). The description-experience gap in risky choice. *Trends in Cognitive Sciences*, *13*(12), 517–523.

Hey, J. D., & Orme, C. (1994). Investigating Generalizations of Expected Utility Theory Using Experimental Data. *Econometrica*, *62*(6), 1291-1326.

Iturbe-Ormaetxe, I., Ponti, G., Tomás, J., & Ubeda, L. (2011). Framing effects in public goods: Prospect Theory and experimental evidence. *Games and Economic Behavior*, *72*(2), 439–447.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291.

Knight, F. H. (1921). *Risk, uncertainty and profit*. Boston, MA.: Hart, Schaffner & Marx; Houghton Mifflin Company.

Kunreuther, H., Silvasi, G., Bradlow, E. T., & Small, D. (2009). Bayesian analysis of deterministic and stochastic prisoner's dilemma games. *Judgment and Decision Making*, *4*(5), 363–384.

Leland, J. W. (1994). Generalized similarity judgments: An alternative explanation for choice anomalies. *Journal of Risk and Uncertainty*, *9*(2), 151–172.

Levati, V. M., Morone, A., & Fiore, A. (2009). Voluntary contributions with imperfect information: An experimental study. *Public Choice*, (138), 199–216.

Ludvig, E. A., & Spetch, M. L. (2011). Of Black Swans and Tossed Coins: Is the Description-Experience Gap in Risky Choice Limited to Rare Events? *PLoS ONE*, *6*(6), 20–26.

McCarter, M. W., Rockmann, K. W., & Northcraft, G. B. (2010). Is it even worth it? The effect of loss prospects in the outcome distribution of a public goods dilemma. *Organizational Behavior and Human Decision Processes*, *111*(1), 1-12.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.

Rakow, T., Demes, K. A., & Newell, B. R. (2008). Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice. *Organizational Behavior and Human Decision Processes*, *106*(2), 168–179.

Rakow, T., & Newell, B. R. (2010). Degrees of uncertainty: An overview and framework for future research on experience-based choice. *Journal of Behavioral Decision Making*, *23*, 1–14.

Ridley, M. (1996). *The origins of virtue*. New York: Penguin books.

Rubinstein, A. (1988). Similarity and decision-making under risk (is there a utility theory resolution to the Allais paradox?). *Journal of Economic Theory*, *46*(1), 145–153.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*(2), 129–138.

Smith, V. L. (2003). Constructivist and Ecological Rationality in Economics. *American Economic Review*, *93*(3), 465–508.

Smith, V. L. (2008). *Rationality in Economics: Constructivist and Ecological Forms* (1st ed.). Cambridge University Press.

Sunstein, C. R. (2003). Terrorism and probability neglect. *Journal of Risk and Uncertainty*, *26*(2), 121–136.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297-323.

Ungemach, C., Chater, N., & Stewart, N. (2009). Are Probabilities Overweighted or Underweighted When Rare Outcomes Are Experienced (Rarely)? *Psychological Science*, *20*(4), 473–479.

van Dijk, E., Wit, A., Wilke, H. A. M., & Budescu, D. V. (2004). What we know (and do not know) about the effects of uncertainty on behavior in social dilemmas. In R. Suleiman, D. V. Budescu, I. Fischer, & D. M. Messick (Eds.), *Contemporary psychological research on social dilemmas* (pp. 315 – 331). Cambridge University Press.

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge University Press.

Appendix

A1. Further statistics

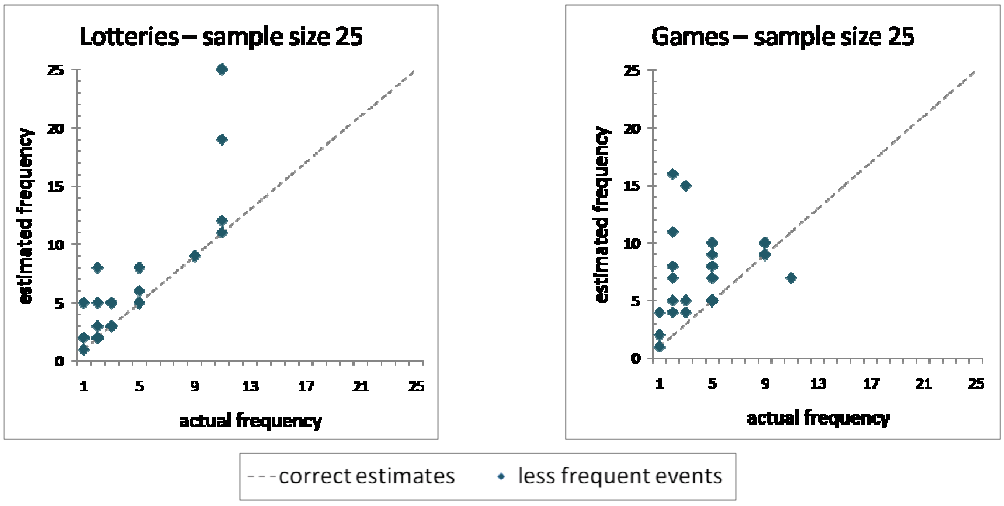

FIGURE A1. FREQUENCY ESTIMATES OF LESS LIKELY EVENTS IN THE LAST ROUND IN LOTTERIES AND GAMES.

TABLE A1. PERCENTAGE OF PEOPLE INVESTING IF 2ND HALF OF THE DRAWN SAMPLE IS LARGER THAN 1ST HALF (N = 256 IN LOTTERIES AND 256 IN STOCHASTIC PGS).

LOTTERY		Percentage of people	
		not investing	investing
Expected outcome larger in 2nd half	no	20%	30%
	yes	14%	36%

STOCHASTIC PG		Number of people	
		not investing	investing
Expected outcome larger in 2nd half	no	29%	18%
	yes	30%	22%

A2. Questions about the primary motives whether to invest in the stochastic PGs

For the cases where participants decided *not to invest* in the common fund, they indicate their main reasoning by ticking one of the following options:

1) probability - "The probability to get a low value for A seemed to high to me, independent of whether the other contributes.

2) value - "The two values for A were not high enough for me, independent of whether the other contributes.

3) conditional non-cooperation - "I thought that the other participant would not contribute."

4) social uncertainty - "I did not know what the other would do. Therefore I did not contribute."

5) greed/opportunism - "I thought that the other participant would contribute and this way I could profit without contributing myself

6) "None of the reasons applies."

7) "I always contributed."

For the cases where participants decided *to invest* in the common fund, they indicate their main reasoning by ticking one of the following options:

1) probability - "The probability to get the high value for A seemed high enough to me, independent of whether the other contributes.

2) value - "The two values for A were high enough for me, independent of whether the other contributes.

3) conditional cooperation - "I thought that the other participant would contribute. 4) social uncertainty – "I did not know what the other would do. Therefore I contributed."

5) moral obligation – "I thought to contribute to the project is the morally right thing to do independent of what the other does."

6) "None of the reason applies.

7) "I never contributed."

TABLE A2. COMBINATIONS OF REASONS FOR INVESTING AND NOT INVESTING

DESCRIPTIOIN

		Contribution							
		Probability	Value	Conditional cooperation	Social uncertainty	Moral reason	Other reason	Never contributed	**Sum**
No contribution	Probability	41%	6%	6%	0%	3%	0%	0%	**56%**
	Value	0%	0%	0%	0%	0%	0%	0%	**0%**
	Conditional cooperation	0%	0%	3%	0%	3%	0%	0%	**6%**
	Social uncertainty	6%	0%	3%	0%	0%	0%	3%	**13%**
	Opportunism	13%	0%	0%	0%	0%	0%	3%	**16%**
	Other reason	0%	0%	3%	0%	3%	0%	0%	**6%**
	Always contributed	3%	0%	0%	0%	0%	0%	0%	**3%**
	Sum	**63%**	**6%**	**16%**	**0%**	**9%**	**0%**	**6%**	

EXPERIENCE

		Contribution							
		Probability	Value	Conditional cooperation	Social uncertainty	Moral reason	Other reason	Never contributed	**Sum**
No contribution	Probability	19%	6%	19%	0%	3%	0%	0%	**47%**
	Value	6%	6%	6%	0%	0%	0%	0%	**19%**
	Conditional cooperation	0%	3%	0%	0%	0%	0%	3%	**6%**
	Social uncertainty	3%	0%	3%	0%	0%	0%	0%	**6%**
	Opportunism	3%	6%	0%	0%	3%	0%	3%	**16%**
	Other reason	0%	0%	3%	0%	0%	0%	0%	**3%**
	Always contributed	0%	0%	3%	0%	0%	0%	0%	**3%**
	Sum	**31%**	**22%**	**34%**	**0%**	**6%**	**0%**	**6%**	

Note. 'Probability' refers to probability being the most important reason, 'value' refers to the value of the multiplier, 'cond. cooperation' refers to whether participants expected the other to cooperate or not, 'social uncertainty' indicates whether being not sure about what the other drives behavior, 'opportunism' indicates whether the player expected the other to cooperate and wanted to use this for her on benefits, 'other reason' indicates that some other reason than the those stated drive the decision, and 'always invested' indicates when participants always invested or did not invest.

A3. Instructions

[Anything in square brackets was not shown to the participants but serves to lead the reader through the instructions. The instructions pertain to the experience condition. At any point where participants are asked to 'draw' this is replaced in such a way that participants are told that they are shown probability statements.]

Dear Participant,
welcome and thank you very much for participating in this experiment!

For arriving on time you receive a payoff of 2.50€. In addition to this, you can earn more money through your decisions in this experiment. Therefore, please read the following instructions carefully. In the experiment you will make several decisions. At the end of the experiment the instructor will determine one round of the experiment at random which will account for your payoff.

Your total payoff from this experiment consists of the amount that you gained in the determined round plus 2.50€ for arriving on time. You will receive your total payoff in cash at the end of this experiment. The payment will happen privately. None of the other participants will know the amount of your payoff.

During the experiment, all amounts will be given in ECU (Experimental Currency Units). **Thereby 1 ECU = 1 Euro.**

Please, stay calm during the experiment and switch off your mobile phones! It is important that you do not take any notes or talk to other participants during the experiment. Communication among participants will automatically lead to a termination of the experiment in which case no participant will receive a payoff from the experiment. If case you have any questions or comments during the experiment, please raise your hand. One of the instructors will then come to your place and answer your question.

[Instructions lotteries]

Instructions

This part of the experiment consists of 8 rounds.

At the beginning of each round, you will receive 10 ECU. Please decide in each round between 2 options:

Option A: **The 10 ECU will be multiplied by 1. You will keep 10 ECU.**

Option B: **The 10 ECU will be multiplied by one of two possible values „A“, one which is high and one which is lower.** Depending on which value of A will be actually realized, the 10 ECU will be increased or decreased.

You will learn about the likelihood of A attaining the high or the lower value by testing how often each of these values will occur. Each time you press on the button „draw“, one of the two values of A will be drawn at random from a card deck and displayed on the screen. The more often a value occurs, the more likely it is.

The program will ask you to draw 25 times. Afterwards you will have to decide for option A or option B. Then, another round will follow.

Please note: You will be informed about the outcome of your decision only at the end of the experiment.

Example

You receive 10 ECU for this round. You have to decide between 2 options:

Option A: You keep the 10 ECU.

Option B: The 10 ECU will be multiplied with A. In this round A equals **0,5 with 10 % <u>or</u> 1,7 with 90%.**

You decide for option B. You will receive either 10 x 0.5 = 5 ECU with 10% or 10 x 1.7 = 17 ECU with 90% probability.

Your final payoff

At the end of the experiment the instructor will randomly determine one round for your final payoff. Then, he will determine one of the two values for A according to its likelihood, thus determining which value of A will be actually realized. The payoff that you gained in this round through your decision will be paid out to you together with the 2.50€ for arriving on time, at the end of the experiment.

The Experiment will now begin. From now on, all decisions that you will make will be relevant for your payoff. Please, remain quiet and seated during the whole experiment and do not talk to each other until you have left the room.

In case you have questions, please raise your hand!

[Instructions Games]

Instructions

Formation of groups

This part of the experiment consists of 10 rounds. In each round you will be part of a dyad interacting with another participant. You will not receive any information on who the other person in your group is; neither during nor after the experiment. For each round new groups will be formed such that no participant will be ever matched with the same counterpart again.

Your task

At the beginning of each round, each participant will receive 10 ECU. Both group members decide independently if they want to contribute 10 ECU to a project. Your payoff depends on your own decision and on the decision of the other group member in the respective round.

Option A: You do not contribute to the project and keep the 10 ECU.
Option B: You contribute to the project. All amounts will be multiplied by a value A. This determines the payoff from the project.

The payoff from the project will be equally split among the two group members; i.e., each group member receives half of the payoff, independent of whether or not they contributed to the project.

$$\textbf{Your part of the project payoff} = \frac{A * (\text{your contribution} + \text{contribution of the other group member})}{2}$$

Then, a new round begins, in which you will be matched with another participant.

Your payoff per round:

Your payoff in each round consists of two parts:

- The amount of ECU, which you do not contribute
- Your part of the project payoff

Your payoff per round = amount of ECU, which you do not contribute + your part of the project

Example 1a:

A has a value of 1.7. You and your counterpart both contribute 10 ECU. The sum of these amounts is 2 x 10 ECU = 20 ECU. The total project payoff is 20 ECU x 1.7 = 34 ECU.

Each of you receives an equal share of this payoff; i.e., 34 ECU / 2 = 17 ECU.

Your total payoff from this round and the total payoff of your counterpart are both 17 ECU.

Example 1b:

A has a value of 1.7. Only you contribute 10 ECU to the project, but your counterpart does not. The sum of these amounts is 10 ECU. The total project payoff is 10 ECU x 1.7 = 17 ECU.

Each of you receives the same part of this payoff; i.e., 17 ECU / 2 = 8.50 ECU.

Your total payoff from this round is 8.50 ECU.

The total payoff of your counterpart is 8.50 ECU from the project plus 10 ECU which he kept for himself = 18.50 ECU.

Amount of the project payoff per round

The amount of the project payoff depends on the value of A. In the last two rounds, A has a fixed value. In each round, there are two possible values for A, a high value and a lower value. Depending on the value that is realized for A, the project payoff will be higher or lower. You will learn about the likelihood of A attaining the high or the lower value by testing how often each of these values will occur.

Each time you press on the button „draw“, one of the two values of A will be drawn at random from a card deck and displayed on the screen. The more often a value occurs, the more likely it is.

The program will ask you to draw 25 times. Afterwards you will have to decide whether or not to contribute your 10 ECU to the project. Then, another round will follow.

Please note: You will be informed about the outcome of your decision only at the end of the experiment.

Example 2:

You receive 10 ECU for this round. You and the other group member decide independently from each other, whether or not to contribute your 10 ECU to the project. The project payoff depends on the sum of your contributions and the value of the multiplier A. In this round, A can take one of two values with following probabilities:

0.8 with 6% <u>or</u> 1.9 with 94%

You and your counterpart both contribute 10 ECU; i.e., in total 20 ECU.

The project payoff either equals 20 x 0.8 = 16 ECU with 6% or 20 x 1.9 = 38 ECU with 94% probability.

Each of you receives an equal share of the project payoff; i.e., either

16 ECU/2 = **8 ECU with 6%** <u>or</u> 38 ECU / 2 = **19 ECU with 94%**

Your final payoff

At the end of the experiment the instructor will randomly determine one round for your final payoff. Then, he will determine one of the two values for A according to its likelihood, thus determining which value of A will be actually realized. The payoff that you gained in this round through your decision will be paid out to you together with the 2.50€ for arriving on time, at the end of the experiment. Please, answer the following comprehension questions, before the experiment starts. Thereby we want to assure that all participants understand the rules of the experiment fully and correctly.

Please, remain quiet and seated during the whole experiment and do not talk to each other until you have left the room. In case you have questions, please raise your hand!

Step-price heuristic: Frugal information use in the used car market

FLORIAN ARTINGER [a,*]

[a] Max Planck Institute for Human Development

Prices are central for coordination in markets. Yet, little is known about the pricing strategies firms employ when faced with an uncertain demand and whether they make full use of the available information. To address these issues, we use interviews with 55 car dealers and pricing data of 748 dealers from an online used car market for two types of cars. Interviewees outlined the pricing strategy they use and evaluated the relevance of nine theories on price dispersion. We find that more than half of all dealers use the 'step-price' heuristic: starting out with a high price, they lower it in fixed time intervals until the car sells. About one quarter of dealers even only consult the online market once for information: when setting the initial price. Price changes at fixed time intervals and frugal information use lead to systematic price rigidity and dispersion as dealers ignore potential changes in the market. Central factors for the use of such a pricing strategy are search for the best price in the face of an uncertain demand and the quality uncertainty that consumers harbor.

Keywords: Sticky Prices; Price Dispersion; Bounded Rationality; Search; Online Market

JEL: D22, D83, L81

* *Corresponding author:* Florian Artinger, Max Planck Institute for Human Development, Center for Adaptive Behavior and Cognition, Lentzeallee 94, 14195 Berlin, Germany, Phone: +49 30 82406 699, artinger@mpib-berlin.mpg.de

"Price setting involves an enormous burden of information gathering and computation that precludes the use of any but simple rules of thumb as guiding principles." (Simon 1962, p. 10)

Introduction

At the largest BMW dealership in Germany that sells used cars. After some time searching, the author discovers a used car that seems just like another car the dealership sells, however, at a much lower price. After some inquiring, the salesman explains that the pricing strategy of the dealership is to keep the price constant for a fixed period, if the car does not sell, the price is reduced. This is repeated until the car sells. The cheaper of the two cars has been simply on the lot for more time. This seems a peculiar pricing strategy which leads to the apparent paradox that two equal cars are priced quite differently at the same dealership. The 'step-price' heuristic the salesman described neglects readily available information inviting sticky prices due to infrequent updating and price dispersion as identical products can deviate in prices. Would such a pricing strategy be widely adopted in the used car market?

A standard assumption in economics is that a perfectly rational profit maximizing firm uses all the available information. Yet, there is little evidence from naturally occurring markets whether firms make full use of the information in setting prices, and how this relates to economic theory (Ellison, 2006). An exception is Fabiani et al. (2007) who use interviews with 11.000 managers in the Euro area. They find that two thirds of firms employ what they refer to as suboptimal pricing rules that only use some of the available information. However, the emergent order in a market does not necessarily hinge on perfectly rational agents who compute a best response from all available information. Gode & Sunder (1993) employ 'zero' intelligence traders in a double auction who randomly choose among bids that do not result in a loss. The agents achieve most of the possible social gains from trade. This indicates that an important element is the institution in which agents operate.

The institution, or environment, in which used car dealers are set, is characterized by an ambivalent situation: the internet provides a high degree of market transparency; at the same time, dealers face a number of uncertainties:

i. how much information does a customer gather;
ii. what is the maximum willingness to pay (WTP) of a customer;
iii. is a customer likely to behave strategically waiting until a price might be reduced;
iv. what is the frequency of customers for a given combination of information (i), WTP (ii), strategic behavior (iii);
v. how do competitors respond to the prices a firm posts;
vi. how do basic economic variables develop.

In line with Simon's (1962) introductory quote, such a setup might preclude the gathering of all relevant information and computation for price setting. This is also reflected in the distinction that Savage (1954) makes between small and large worlds: Bayesian decision theory, of which Savage is the founding father, can only be applied in small worlds where all alternatives and probabilities are known. In a large world, the decision maker lacks complete knowledge and is faced with uncertainty where optimization is out of bounds (Binmore, 2007). The use of only some of the available information is an essential element of a fast and frugal heuristic, an evolved decision strategy that is well adapted to a given environment (Gigerenzer, Todd, & the ABC Research Group, 1999).[1] The extent to which the environment and the strategy fit has been referred to as ecological rationality (Gigerenzer et al. 1999; Smith, 2008). Step-price, due to its use of only a limited amount of information and since it likely evolved to be adaptive in the used car market, would constitute a fast and frugal heuristic.

[1] Fast and frugal heuristics are distinct from the heuristics and bias school (Tversky & Kahneman, 1974). Both agree that decision makers frequently rely on heuristics, i.e., where decisions are based on a limited amount of information. Heuristics and biases emphasize that the decision strategies that people use frequently fall short of a normative standard, usually a perfectly rational profit maximizer of the as-if tradition by Friedman (1953). It mainly points to situations where there exists a mismatch between decision strategy and environment for instance due to the lack of an opportunity to learn. For a discussion on both schools see Kelman (2011).

Taking into account the full complexity of pricing in the face of a dynamic, uncertain, demand where firms have some price setting power has been recognized in the management literature to be mathematically intractable (Aviv & Pazgal, 2005; Lovejoy, 1991). The complexity originates from the high dimensionality of the decision problem in combination with the strategic interaction of firms and consumers. In response, the pricing algorithms developed started to test the effects if certain information is ignored. For instance, Aviv & Pazgal (2005) develop an approximate model which disregards the impact of the current price a firm sets on the decision of other firms. They find that their pricing algorithm performs surprisingly well compared to a benchmark model computed under discrete and finite state space. Similarly, Mersereau & Zhang (2011) develop a robust pricing strategy that requires no knowledge of the frequency of strategic customers and find that such a policy does remarkably well compared to a model with deterministic demand where sellers have full information. A more radical approach is described by Selten (1998) in the aspiration adaptation theory: a decision maker searches for new information and sequentially adapts the targets. The decision maker does not maximize, not even with regards to a subset of the available information, unlike the algorithms of Aviv & Pazgal (2005) and Mersereau & Zhang (2011).

If step-price is in fact a widely adopted strategy because it is well adapted, this also should give rise to the observation of price stickiness and dispersion. Both phenomena have been viewed as economic regularities persisting in markets. Price stickiness (or rigidity) implies that prices react only slowly to changes in the market or the cost structure of firms. Price dispersion stands in contrast to the 'law of one price'. Economic textbooks state that in the case of price competition (Bertrand) with a homogenous product and absent any search costs all firms charge the same price. Empirical evidence suggests that this does not hold, not even in a highly transparent market where search costs are minimized (e.g., Brynjolfsson & Smith, 2000, for a review see Baye, Morgan, & Scholten, 2006). The literature on sticky prices and dispersion has largely developed separately. However, if firms employ a step-price strategy and are not perfectly synchronized in their price setting, sticky prices necessarily lead to price dispersion.

A number of theories on price stickiness have recently found empirical support via the use of interview studies with managers. However, none of these theories addresses the effect of an uncertain demand. This paper links the empirical observations of the pricing strategies of the used car market to current theories on price stickiness and dispersion especially focusing on the effect of uncertain demand.

We use interviews conducted at 55 dealerships and data from an online used car market where the pricing decisions of 748 dealers in Germany for two car models for a period of nine months were tracked. In 2008, 6.11 million used cars were sold compared to 3.09 million new cars in Germany (Michel, 2009), making the used car market an important source of revenue for dealers and an ideal field to study information use in pricing. Our results show that more than half of all dealers employ the step-price heuristic. In combination with the interviews the online data suggest that dealers persistently ignore information that is readily available to them. Important elements that drive this behavior is search for the best price but also quality uncertainty on the side of the consumer. Our results help to better understand why price stickiness and dispersion persist in markets and can serve future theory development.

In the next section, the theoretical background is outlined. This is followed by the results where first the interviews are discussed. This provides important base on which the subsequent analysis of the online data rests. A discussion concludes.

Theoretical background

Assuming that firms use all available information when setting prices, economic theory has spawn numerous models on price stickiness and price dispersion. The most important theoretical explanations are summarized below.

Sticky prices

A number of theories have been developed addressing why prices are sticky (for a review see Blinder, Canetti, Lebow, & Rudd, 1998); however, conventional econometric testing of competing models has been out of bounds. It has been difficult to find a metric to measure the speed at which the market clearing price

adjusts which could serve as a benchmark for price stickiness. A second aspect why empirical testing has been difficult is that many of the most prominent theories rely on unobservable variables (Blinder et al. 1998). As a solution to this conundrum, interviews with managers that set prices have been conducted in order to evaluate which theory describes actual pricing decisions well (Amirault, Kwan, Wilkinson, & Canada, 2006; Apel, Friberg, & Hallsten, 2005; Blinder, 1991; Blinder et al. 1998; Fabiani et al. 2007; Hall, Walsh, & Yates, 2000).

To evaluate whether theories on price stickiness that have been tested to date are relevant for the used car market, we selected eight out of twelve theories investigated by Blinder et al. (1998). Four theories were dropped as interviewees in previous studies did not regard these as relevant to price setting. These theories also do not relate to the used car market. The following summarizes the eight theories on price stickiness presented to interviewees:

1. Cost based pricing. The cost of material and labor are essential for price calculations. The assumption here is that even in the face of a changing market if costs stay constant, prices do not change instantly (e.g., Bils, 1987).
2. Implicit contracts. Firms would like to build long term relationships with customers. One tool in doing so is to change prices as little as possible. Even if there are changes in the market or cost structure, firms might not instantaneously respond with a priced change. This is modeled by Okun (1981).
3. Explicit contracts. Firms have contractual agreements with customers which determine prices in advance of a transaction (e.g., Fischer (1977) who first introduced this idea for the case of wages).
4. Co-ordination failure. If a firm would raise the price, customers buy at competitors that maintain the cheaper price; likewise if a firm reduces its price, competing firms need to reduce prices to attract customers. If there is no coordinating mechanism which allows all firms to move together in case of changes in costs or the market, prices remain constant (e.g., Cooper & John, 1988).
5. Non-price competition. Prices are not necessarily the only element that facilitates market clearance. Instead of adjusting the price, firms can for instance adjust service keeping the price constant (e.g., Carlton, 1986; Maccini, 1973).

6. Price points. Blinder et al. (1998) refer to the marketing and psychology folklore whereby prices are set to psychologically attractive numbers. Levy, Lee, Chen, Kauffman, and Bergen (2011) show that 9 is the most frequent price ending for a number of retail products in a brick-and-mortar store and on the internet. The most frequent price change is one where the terminal digit is 9. If prices have to 'jump' to the next psychologically attractive digit, they might be less sensitive to changes in costs or the market.

7. Judging quality by price. Firms do not reduce a price in a slack market out of fear that this is interpreted as a reduction in product quality (e.g., Allen, 1988).

8. Menu costs. Changing prices in itself is costly, due to costs of advertizing these anew, printing new price tags etc. Therefore, prices do not adjust perfectly to changes in market conditions (e.g., Akerlof & Yellen, 1985; Mankiw, 1985).

TABLE 1. RANKING OF EIGHT THEORIES ON PRICE STICKINESS AND MEAN RANK.

	Mean rank	**US** (Blinder et al. 1998)	**UK** (Hall et al. 1997)	**CA** (Amirault et al. 2004)	**SE** (Apel et al. 2005)	**Euro area** (Fabiani et al. 2007)
Cost-based pricing	2.0	2	2	1	2	3
Implicit contracts	2.6	4	5	2	1	1
Explicit contracts	2.8	5	1	3	3	2
Co-ordination failure	3.4	1	3	5	4	4
Non-price competition	5.5	3	8	4		7
Price point	7.3	8	4		7	10
Judging quality by price	9.0	12	10			5
Menu costs	9.2	6	11	10	11	8

These theories have been tested in five interview studies across a wide range of industries using interviews with several thousand managers. The results are shown in table 1. Cost-based pricing, implicit and explicit contracts, and coordination failure are always under the top five rated. Yet, these theories do not address in any way how firms might respond to an uncertain demand.

Price dispersion

Theories on price dispersion suggest that it can occur because of asymmetric information (Baye & Morgan, 2001; Varian, 1980). If some customers are better informed than others, it can pay off for stores to randomize prices, i.e., prices

move in such a fashion that customers cannot infer which store is consistently more expensive. There is empirical evidence from consumption goods that firms indeed randomize prices (Lach, 2002).

Asymmetric information is particularly relevant in the used car market due to the complexity of the product. This is reinforced as consumers frequently cannot ensure that the car does not have any hidden defects. The resulting quality uncertainty potentially invites dealers to randomize prices as consumers cannot precisely link price with quality. However, if the issue of quality uncertainty is widespread, this can result in the adoption of institutions like brand name and reputation to distinguish potentially fraudulent dealers from those that offer higher quality products (Heal, 1976).

If price changes are not perfectly synchronized, as it is likely with a step-price strategy, this can mimic some degree of randomization. If this would be a principle mechanism that drives the use of step-price, no car dealerships should be systematically more expensive than another dealership. Price randomization and the impact of reputation are tested using the online data.

Search

None of the above theories on price stickiness and dispersion addresses what firms might do in the case they face an uncertain demand. Lazear (1986) suggested that periodic price changes constitute an effective tool to learn about an uncertain demand. In this model, which lays the foundation on which Aviv & Pazgal (2005) and Mersereau & Zhang (2011) build their algorithms, a firm has price setting power and faces an uncertain but uniform demand. Initially, a high price is set to address customers with a high WTP. Subsequently the price is lowered. However, such a policy is put into question if customers act strategically as noted by Coase (1972, see for a formal model Bulow, 1982). In response, Conlisk, Gerstner, & Sobel (1984) develop a model of a monopolist where customers differ in their WTP and their preference to purchase the good immediately or after a delay. Customers who enter have the same distribution in terms of WTP and strategic orientation as those in the market at period 1. Customers who buy the product leave the market. After a number of periods, a sufficiently large number of customers is left with a high strategic orientation

aiming to buy the product for a reduced price where it pays-off for the firm to lower the price. In order to evaluate the relevance of learning about an uncertain demand, the proposition is given to interviewees to rate as question 9.

Price stickiness and dispersion have also been modeled as the result of a learning process. Rothschild (1974) shows how a sampling error can lead to price dispersion as the samples from which dealers learn might differ and there is a tradeoff between learning about the true WTP and profits. Balvers & Cosimano (1990) suggest that large price adjustments would introduce noise that makes it difficult to learn. An optimal tradeoff between learning about the price and adjusting entails small price movements that cause price rigidity. In both models, concurrent updating is taking place thereby making use of all the available information.

Analogously to any of the previous theories, searching for the best price as implied by the models above employs some form of maximization. Decision makers might in fact agree that search is an important aspect, yet, the decision process likely differs. Therefore it is important to complement the interview data with the empirical observations from the online market to get an accurate understanding of the pricing strategies applied and about their ecological rationality.

Hypotheses

In face an uncertain demand and an environment that is marked by complex strategic interaction, the step-price heuristic constitutes a robust decision process which ignores readily available information.

HYPOTHESIS 1: DEALERS EMPLOY THE STEP-PRICE HEURISTIC IN THE USED CAR MARKET.

A number of theories on price stickiness and dispersion have been developed. However, none of these theories stress the importance of a search process on the side of the firm for the best price. In the environment that used car dealers operate in this is a likely behind the use of the step-price strategy.

HYPOTHESIS 2: DEALERS EVALUATE SEARCH FOR THE BEST PRICE AS A THEORY THAT IS RELEVANT TO THEIR PRICING DECISIONS.

Results

Interviews

The interview contains three parts (see Appendix A1 for the complete interview manuscript). The first part consists of questions regarding the specific characteristics of a dealership and the background of the interviewee. In the second part, pricing strategies are investigated. In the third part, we inquire about the relevance of nine theories on pricing and how these relate to the practice of the dealers. In this last part of the interview, the order of questions, i.e., theories, is randomized to prevent order effects.

The interviewees were recruited from a list of 902 used car dealers listed on Autoscout.de and selling used BMWs in Germany. In order to minimize noise in the data we focused on dealerships that operate in a more competitive environment. To measure competitiveness, we used the city and ZIP code of dealerships to calculate the total number of used car dealers in a given combination of city and ZIP code. We rank this data from highest to lowest density of the number of dealerships. All dealerships up to rank 450 were sent an invitation by mail to participate in the interview. Overall, 55 dealerships agreed to be interviewed. 30 interviews were conducted in person and a further 25 via telephone. Interviewees were randomly allocated to one of the two ways of conducting the interview. The responses were not different depending on the method used to gather the data (testing the pricing strategies that dealers employ, these are not statistically different whether interviews in person or by phone, $\chi^2(6) = 7.95, p = .21$, the same holds for the evaluation of the theories $\chi^2(9) = 2.38, p = .98$). The dealerships were geographically dispersed across Germany, resulting in a distance travelled for the interviews that were conducted in person of close to 3.000 km.

The interviews lasted about 20 minutes and were recorded on audio tape and later transcribed. These were given to two independent raters to categorize

answers. The inter-reliability of raters is 96%; where differences existed raters discussed these and agreed on a common rating.

Background of interviewees

The interviewees have a mean experience in pricing used cars of 17.8 (SD = 11.1) years; 11% are owners, 60% general managers, and 29% head of sales. The 55 dealerships at which the interviews were conducted offer in the mean 101 (SD = 125) used cars. Used cars make up 68% (SD = 20) of all cars on sale at the mean dealership, the remaining share are new cars. Of all used cars at the mean dealership, 76% (SD = 32) are BMWs. The mean dealership has about 28% (SD = 22) of returning customers, i.e., customers that have bought at least their last car at the dealership. Embedded in a larger network with a centralized pricing system for used cars are 12 of the 55 dealerships. In such a network with centralized pricing there is a mean of 6.0 (SD = 5.9) dealerships connected.

Pricing strategies

Using a mix of open and closed questions in the second part of the interview we investigate which pricing strategies dealers use. Close questions allow to precisely enquire about the mechanisms dealers use and to generate more readily comparability across interviews. At the same time, using open questions no details are lost that would have not been captured if one only employs closed questions.

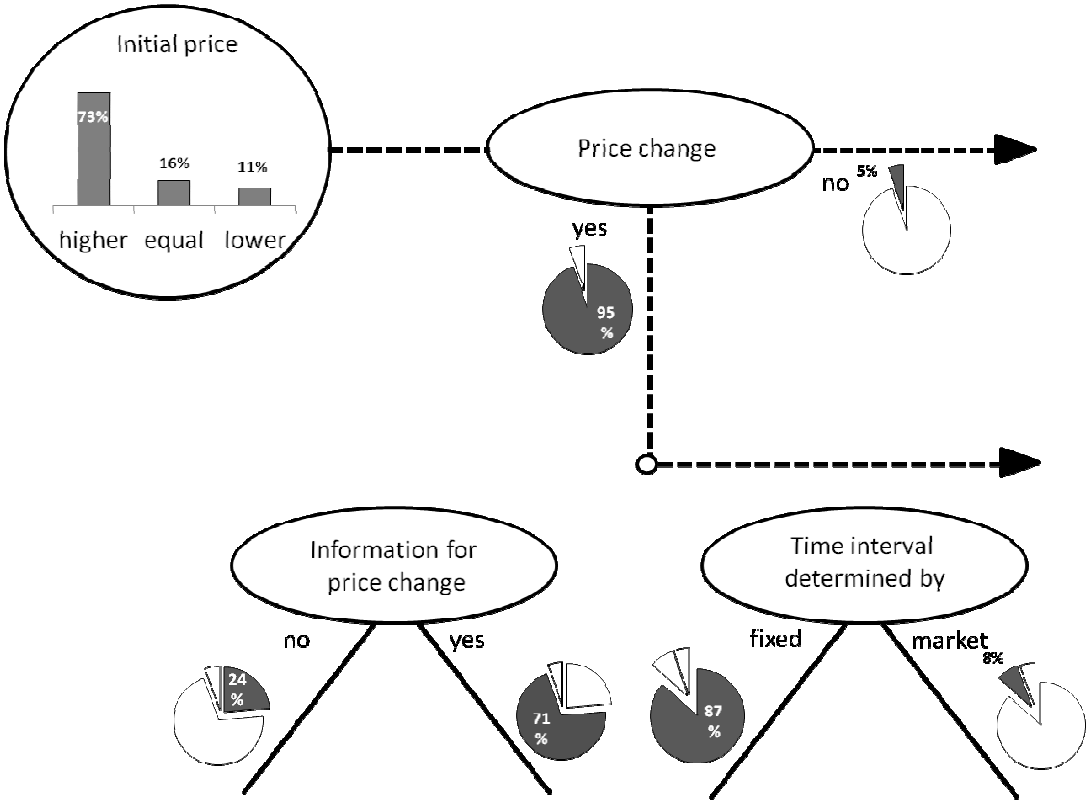

FIGURE 1.USE OF PRICING STRATEGY COMPONENTS (N = 55).

The results on the initial price and price change including the information use are illustrated in figure 1. 73% of the interviewees set the initial price higher than the price of a comparable car of a competitor that is currently in the market; 11% underbid and 16% match the price. Price changes are regularly undertaken by 95% of interviewees. Overall, 87% of interviewees employ fixed time intervals at which they would consider a price change; the remaining 8% change prices if they observe shifts in the market or the cost structure. 54% of interviewees consider a price change for a used car every 30 days. Further 33% either have a longer fixed time interval than 30 days before they consider a price change, or initially wait for 30 days followed by shorter periods at which they consider to change the price of a car.

When setting the initial price of a used car the mean number of information sources an interviewee consults is 2.0 (SD = .70). 98% of the interviewed dealerships consult one or more of the online markets to compare going prices of similar cars. Further 64% use a market study that lists the value of a used car

given a number of conditions.[2] When collecting information, 53% focus on Germany as a target market, 27% on the local market (100 km in diameter around the dealership), 12% on the regional market (200 km in diameter), and 8% on Europe.

The total percent of those who regularly change prices account for 95% of all interviewees. Of these 71% consult the online market or market surveys again. However, 24% already determine with the initial price by how much a price will be reduced if the car does not sell. They do not consult any further information when prices are changed.

Considering the pricing strategies that are most frequently applied, 45% start with an initially high price and undertake price changes at fixed time intervals. Information from the online markets or market survey inform these changes. The second most prominent pricing strategy used by 24% of interviewees also entails starting with a high price, reducing the price at fixed intervals but, different to the first strategy, by an amount determined as soon as the car enters the lot. Both strategies together reflect the use of a step-price heuristic, whereby the price is sequentially lowered if no sale occurs after a fixed time period. Thus, step-price accounts in total for 69% of the pricing strategies in the interviews. A further 11% of interviewees start with a price that is similar to that of competitors with price changes at fixed intervals where new information from the market or surveys is used. There are a number of further strategies employed that each individually accounts for less than 5% of all strategies and in sum for 20%.

Do dealers employ multiple pricing strategies? 71% of interviewees state that they only employ one pricing strategy, 29% use some variations but these do not apply to the standard used car sold to consumers but are employed if cars are sold to other dealerships or for special used cars such as old timers. All interviewees indicate that they advertize cars immediately online once a car is on the lot. This implies that the online market does provide a very good source with regards to the current market except for those 10 percent of cars that are not listed. Do interviewees at any one time aim to offer the cheapest car within a set of

[2] These are equivalent to the 'Kelley Blue Book' in the USA.

comparable cars? 58% of interviewees state, that if a car is not sold within at most 180 days they will set the price of a car such that it is the cheapest in the market. However, for 27% of interviewees this is not a good strategy as in their opinion the cheapest car likely constitutes a fraudulent offer. The remaining 9% of interviewees that do implement price changes opted not to respond to the question. In the mean interviewees consider about 40% (SD = 32) of customers to be price sensitive.

Pricing theories

Interviewees responded to nine statements that each reflect a particular pricing theory. Each interviewee graded each theory according to following marks: the theory is

1 = totally unimportant 2 = of minor importance

3 = moderately important 4 = very important

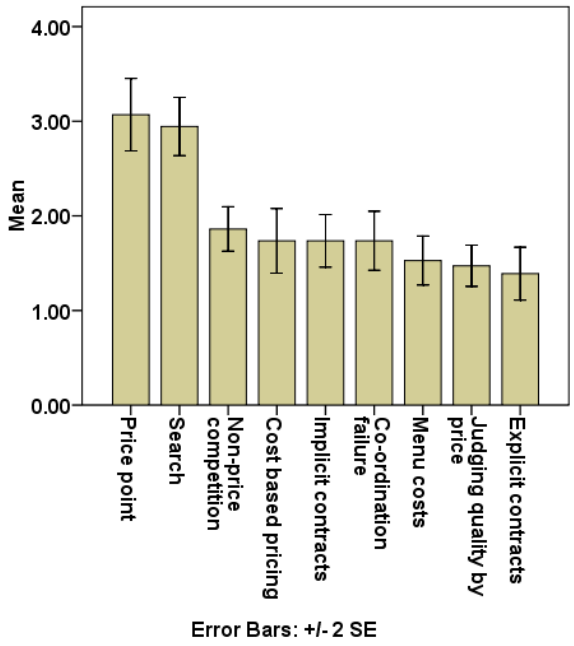

FIGURE 2. MEAN RATINGS OF RELEVANCE OF PRICING THEORIES FOR PRICING OF USED CARS FOR EIGHT THEORIES PREVIOUSLY INVESTIGATED AND SEARCH IN THE FACE OF AN UNCERTAIN DEMAND.

As can be seen in figure 2, only two theories receive in the mean more than 2 points. Blinder et al. (1998) regards this as a critical mark below which a theory bears relatively little relevance for actual pricing. Most important for practitioners

in the used car business are price points which receives a mean of 3.0 (SD = 1.2) suggesting that prices might not change frequently due to 'jumps' to certain, psychologically attractive, numbers. A further theory that is judged as highly relevant is searching for the best price with a mean rating of 2.9 (SD = .9). The evaluation of theories is not different depending on whether dealerships use step-price or not ($\chi^2(9) = 8.92, p = .45$,see table A1 in the appendix for details on the grades).

How do these results compare to the five interview studies previously conducted (Amirault et al. 2006; Apel et al. 2005; Blinder et al. 1998; Fabiani et al. 2007; S. Hall et al. 2000)? The converging evidence across these studies is that the theories on cost-based pricing, implicit and explicit contract, and co-ordination failure drive price stickiness. However, these theories do not feature strongly in the case of the used car market indicating that price stickiness might be driven by different elements depending on the context in which prices are set.

In summary, 69% of dealerships employ step-price, initially starting with a high price and updating prices infrequently in fixed time intervals. 24% of dealerships only use market information once: to set the initial price. The theories on pricing points and on search are rated as highly relevant to pricing by interviewees.

Online market

The online used car market Autoscout.de allows to track the pricing decisions of used car dealers and to analyze how dealer specific characteristics relate to pricing decisions. Autoscout.de covers 78% of the used car market in Germany and allows tracking pricing decisions (Dudenhöffer & Schadowski, 2011).[3] We focus on two types of cars, 320 and 730 from BMW. BMW is one of the largest car manufacturers with about 3 million cars in Germany in use in 2010 (BMW Group, 2011). The online data was collected for nine months starting 8th of December 2010. At the start of the online data collection in December 2010, the car type 320 was the most widely traded BMW on Autoscout.de. The 730 BMW is more

[3] There are two dominating market platforms for used cars in Germany, mobile.de and Autoscout.de. Mobile.de covers 90% of the used cars on offer, however, does not readily permit data collection. Most used car dealers advertize on both websites.

expensive than the 320 and serves to investigate whether dealers that offer a larger number of highly priced cars tend to employ a different pricing strategy. Of the collected data, the 320 BMW sells in the mean for 25,670€ (SD = 9,434), 730 BMW sells in the mean for 53,133€ (SD = 18,819).

The data was collected by searching Autoscout.de bi-weekly for all 320 and 730 BMW. The results pages were downloaded and the data for each car extracted. The results pages contain 12 attributes for each car: price, mileage (in km), the date when the car was first put into service, horse power, type of petrol, type of gears, car body, extra warranty that the car is sold with, ZIP code and city where the dealership is located. We use all offers posted by car dealers in Germany excluding used cars damaged in a car accident. For the 320 overall 352,603 prices are observed. On a single day, there are in the mean 4,415 (SD = 274.9) cars of the 320 type on offer. There are in total 31,021 cars of the 320 type for which prices are recorded. For the 730 BMW there are 37,820 prices observed. On a single day, there are in the mean 470 (SD = 41.1) cars of the 730 type on offer. For the 730 BMW there are 2,731 cars for which prices are recorded. In the mean it takes 47.7 (SD = 31.5) days until a car sells.

A caveat applies to the number of dealerships for which pricing strategies can be observed: the results page of Autoscout.de does not contain the name of the dealership but only the ZIP code and city where the car is sold. Some dealerships have the same ZIP code and city, which implies that a unique attribution of the observed prices to a specific dealer is possible for 748 out of 871 dealers with 24,482 cars from originally 33,752 observed cars.[4] For dealerships that can be uniquely identified we observe in the mean 32.7 (SD = 65.5) cars during the entire period of data collection, the median is 13.0. At the mean dealership, 93.7% (SD = .14) of the observed cars are 320, the rest 730 BMW. The median is 100% 320.

[4] Autoscout.de listed 902 used car dealers selling BMWs from which interviewees were selected. Of these 871 were active during the data collection period.

Pricing strategies

There are three principle components of a pricing strategy for used cars that can be observed in the online data: i) initial price; ii) size of a price change; iii) time span a price is maintained. In a first step, these are analyzed separately.

Components

In order to evaluate the level the initial price and the extent of price changes a reference class is needed. For this we match *for each day* those cars (= prices) that have identical values on following attributes: type of car, year put into service, horse power, type of petrol, type of gears, car body, extra warranty that car is sold with, and mileage differing by a maximum of 10.000 km. In total, for 81.17% of the recorded prices a match can be found. In the mean the number of cars (= prices) that match results in 9.0 (SD = 10.4) cars being grouped together, in the median a group contains 5 cars. This procedure allows determining the minimum price that is posted on any given day for a group of matching cars.

The histogram in figure 3A shows the initial price dispersion. The initial price dispersion is computed using the mean relative difference of the initial price minus the minimum price of the cheapest, matching car. The difference is standardized using the minimum price as a denominator and a mean is computed for each dealer. The bin size is 1% deviation of the initial price from the minimum price. A first inspection of figure 3A indicates that there is a considerable number of dealers who post an initial price of a car that is very competitive. In total, 19.5% of dealers are contained in the first bin. Of these, 16.0% of dealers start in the mean with a price such that they are the cheapest in the market. However, 9.6% of dealers do not have a single car that matches another car resulting in that these dealers cannot have an initial price dispersion larger than 0%. This leaves 6.4% of dealers who start out in the mean with a price that is the lowest in a group

of matching cars. Taking into account the entire sample, the mean dealer posts an initial price of 7.8% (SD = 6.6) above the minimum price.[5]

Does the initial price dispersion decrease the greater the number of cars that match each other, i.e., the greater the competitive pressure?[6] There is a large spread ranging from cars for which no match is found to groups where 111 cars match each other. If competitive pressure reduces initial price dispersion, it would follow that there is a negative correlation between the number of cars that match each other and the difference of the initial price from the minimum price. However, the correlation is in fact positive with r = .40 (p = .001). This suggests that the higher the degree of competitiveness and the larger the number of cars that match each other the greater the initial price dispersion. A likely mechanism that yields this result is that a new car compared to a group of matching cars tends to be priced higher than those already in the market.

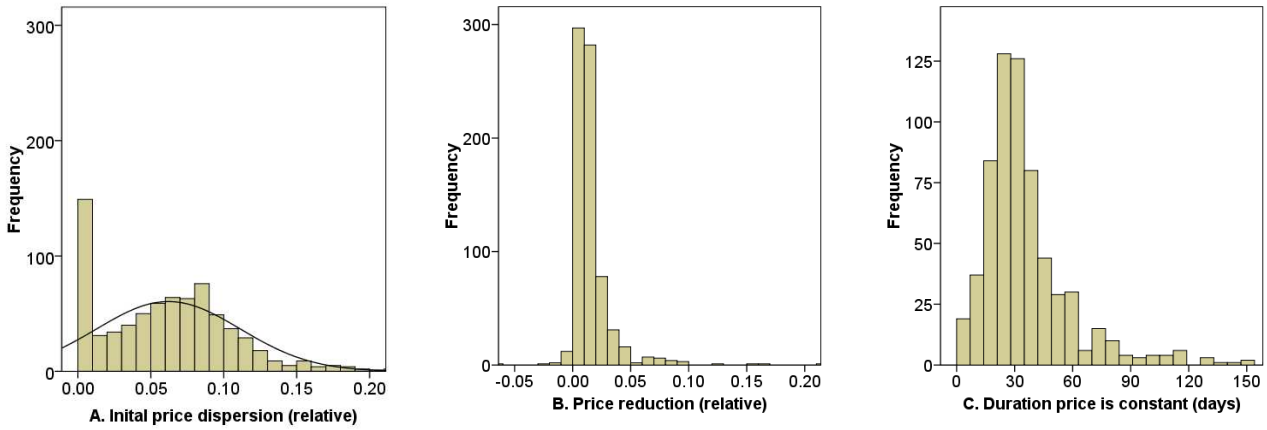

FIGURE 3. HISTOGRAMS USING THE MEAN VALUE PER DEALER FOR EACH OF THE THREE PRICING COMPONENTS.

3A. MEAN RELATIVE DIFFERENCE OF INITIAL PRICE FROM MINIMUM PRICE (N=748). 4B. MEAN RELATIVE PRICE REDUCTION (N=748). 4C. MEAN DURATION BEFORE A PRICE IS CHANGED IN DAYS EXCLUDING THE LAST PERIOD WHEN CAR IS SOLD (N=637).

[5] If one excludes dealers who do not have any car for which another matching car can be found the mean dealer's initial price is 8.13% (SD = 5.00) above the minimum price, the median dealer is 7.54% above the minimum price.

[6] Research shows that thin markets with few buyers and sellers display a substantial price dispersion, which is reduced the thicker a market gets (e.g., Tauchen & Pitts, 1983; Telser & Higinbotham, 1977).

The second component of a pricing strategy observed in the online market is the size of a price change. Figure 3B shows a histogram with the mean price reduction per dealer. A price change that a dealer undertakes is standardized by taking the difference between the old and the new price using the old price as a numerator. The bin size is 1%. In the first bin in figure 3B to the right of 0 there are 39.7% of dealers who change the price in the mean by not more than 1%. Never changing the price of a car is done by 14.84% of all dealers. The mean dealer reduces the price by 1.5% (SD = 2.1).[7]

The third component of pricing strategies is the length of the time interval until a price changes. To make an inference about the pricing strategy of a dealer, one needs to consider the time span until a dealer changes the price of an unsold car. We observe this for 637 of 748 dealers. Figure 3C displays the histogram of the mean days before a price is changed excluding the final period in which the car is sold. Since the data was collected bi-weekly as a conservative measure the data is binned in time spans of 7 days. Only 3.0% of dealers fall into the first bin of changing a price at least once every 7 days. In the mean, a dealer changes the price of a car after 36.2 (SD= 24.0) days.

Differences between dealers

The above analysis highlights that there is a considerable spread contained in each of the three pricing components. In the following a cluster analysis addresses to what extend this spread can be attributed to a number of homogenous subgroups of dealers. Using a multinomial logistic regression, we show how certain characteristics of a dealership influence the choice of a pricing strategy.

The cluster analysis is based on the three pricing components of each dealer. It employs the Ward method with the squared Euclidian distance as a measure for proximity. The goal of the Ward method is to minimize the variance within a cluster. The cluster analysis treats 111 dealers that do not change their price as a separate cluster. The remaining 637 dealers are grouped into three clusters shown in table 2A. The first row in table 2A shows the number of dealers, the second

[7] Only considering those dealers that do use price changes, the mean dealer reduces the price by 1.8% (SD = 1.8).

displays the number of cars that the 748 observed dealers posted online on the 15th of May 2011. This serves as an approximation of the total size of a dealership. The 748 dealers represent in total 55,352 cars which they jointly offer on a single day. For each of the three pricing components, mean, median and F-values are listed. The F-values in table 2A indicate by how much the variance of the component for this cluster is reduced compared to the variance of the complete data set. A value below one indicates a lower variance and hence larger degree of homogeneity than in the complete data set. A value above one indicates greater variance than in the complete data set.

The cluster 'Step-price' contains 56% of dealers selling 71% of cars. The initial price is the highest of all clusters with a mean of 10.4% above the minimum price. Prices are changed in the mean every 35.8 (SD = 15.4) days. This cluster is particularly homogenous; all of the F-values are below one. The second cluster is labeled 'Competitive' as dealers that belong to this cluster tend to compete on price. They start with an initial price that is 2.5% above the minimum price and reduce it in the mean every 21.1 (SD = 9.0) days. The competition on price is also reflected in the mean rank in the group of matching cars. This mean rank is 2.0 (SD = 1.7); the median is 1.5, with 1 indicating that a car is the cheapest within a group of matching cars (see table A2 in the appendix). This cluster is still relatively homogenous. Only the component 'price change' receives a larger variance than in the complete data set. The last two clusters 'Wait – price change' and 'Wait – no price change' are similar: dealers in these clusters either change prices after they have been posted online for an extensive period ('wait – price change') or not at all ('wait – no price change') A lot of the heterogeneity of the data set is captured in the cluster 'Wait – price change' that contains the smallest number of dealers: none of the F-values is below one, especially the variance for 'price change' is considerably larger than in the complete data set.

TABLE 2A. CLUSTER ANALYSIS.

	Complete data		Step price			Competitive			Wait - price change			Wait - no change		
Number of dealers (share)	748 (100%)		423 (56%)			171 (23%)			43 (6%)			111 (15%)		
Number of cars (share)	55,352 (100%)		39,490 (71%)			8,598 (16%)			3,444 (6%)			3,820 (7%)		
Dealers	Mean	SD	Mean	SD	F-value	Mean	SD	F-value	Mean	SD	F-value	Mean	SD	F-value
Initial price from min price	7.7%	6.6%	10.3%	5.8%	0.8	2.5%	2.4%	0.1	8.6%	7.5%	1.3	5.1%	8.0%	1.4
Price change	1.5%	2.1%	1.5%	1.0%	0.2	1.9%	2.8%	1.9	3.7%	4.7%	5.2	0.0%	0.0%	-
Time interval without last period (days)*	36.2	24.0	35.8	15.4	0.4	21.1	9.0	0.1	80.3	73.1	1.6			-

* 'Time interval without last period (days)' for the complete data refers to N = 637 dealers.

TABLE 2B. MULTINOMIAL LOGISTIC REGRESSION WITH STEP-PRICE AS REFERENCE CATEGORY

Reference category: Step-price (N= 423)	Coef.	p	Robust Std. Err.
Competitive (N = 171)			
Number of of 320 and 730	-0.02	0.00	0.00
Proportion of 320	-0.58	0.41	0.70
Competitive setting - number of cars	0.00	0.72	0.00
Mileage	0.00	0.00	0.00
Proportion of cars with extra warranty	-0.23	0.50	0.33
Part of network	-0.20	0.48	0.28
Official partner of BMW	-0.88	0.00	0.23
Constant	-0.05	0.94	0.71
Wait - price change (N = 43)			
Number of 320 and 730	-0.01	0.44	0.02
Proportion of 320	-0.87	0.44	1.12
Competitive setting - number of cars	0.00	0.99	0.00
Mileage	0.00	0.06	0.00
Proportion of cars with extra warranty	-0.42	0.51	0.63
Part of network	-0.38	0.47	0.53
Official partner of BMW	-0.51	0.20	0.40
Constant	-1.42	0.31	1.39
Wait - no price change (N = 111)			
Number of 320 and 730	-0.26	0.00	0.05
Proportion of 320	-1.04	0.11	0.65
Competitive setting - number of cars	0.00	0.63	0.00
Mileage	0.00	0.05	0.00
Proportion of cars with extra warranty	-0.07	0.88	0.47
Part of network	-0.32	0.51	0.49
Official partner of BMW	0.07	0.83	0.34
Constant	1.45	0.06	0.77

N total = 748, Pseudo R^2 = .19, Log likelihood = -666.52

Besides prices, the online market also provides further data that allow the computation of eight characteristics that mark each dealership (see table A2 in the appendix for the mean, median, and SD). These include the (1) number of cars of 320 and 730 BMWs per dealer, (2) the share of 320 BMWs, (3) and the total number of cars advertized online on May 15th 2011. Furthermore, the (4) competitiveness of the environment in which the dealership operates is assessed by using the first two digits of the five-digit German ZIP code which marks an area of 76 km in diameter. Those dealerships that share the first two digits are assumed to share a competitive environment. The total number of cars that a dealership advertized is then used as a proxy to evaluate how many cars in total are offered in such a competitive environment. Further characteristics are the (5) mean mileage of the cars at a dealership as proxy for quality of the cars on offer, (6) the proportion of cars at a dealership that are offered with an extra warranty, (7) whether the dealership is part of a network of car dealers, (8) and the proportion of dealers that are official BMW partners as a signal for reputation.

In order to investigate the influence of these independent variables on the choice of a strategy of a dealer we use a multinomial logistic regression.[8] The dependent variable is the log of the probability with which the dealer belongs to one of two pricing clusters as depicted in equation 1. The reference cluster 'Step-price' is held constant. This results in three regression models as shown in figure 4B, the numerator being either cluster 'Competitive', 'Wait – price change', or 'Wait – no price change'.

$$\log\left[P(clusterX)/P(step\ \ price)\right] = b + \sum_{i=1}^{7} w_i b_i \qquad (1)$$

The only independent variable that significantly distinguishes the probability that a dealer is part of the 'Step-price' cluster across all three models is mileage. However, the coefficient is 0. Considering the mean and median mileage of each cluster (see table A2 in the appendix), this suggests that dealers who employ step-

[8] A correlation analysis shows that total number of cars advertized online (variable 3) correlates with number of cars of 320 and 730 BMW per dealer (variable 1) with r = .75 (p = .001). We keep the number of cars of 320 and 730 BMW per dealer in the regression. This results in seven independent variables as also shown in table 2B.

price tend to sell cars that have less mileage compared to dealers that employ a different strategy. Comparing the clusters 'Step-price' and 'Competitive' we find that the more observations and hence the larger a dealership, the smaller is the likelihood that a dealer is part of the 'Competitive' cluster. Furthermore, if dealers are an official partner of BMW they are less likely to employ a 'Competitive' pricing. Interestingly, the competitiveness of the environment does not influence the probability of whether dealers employ step-price or any of the other three pricing strategies.

In summary, the initial price of a car is often set so that it is not the cheapest in a group of matching cars. In fact, more cars that are matched in a group, the higher is the initial price. Price reduction is used widely but only after some time that prices are kept constant and in relatively small steps compared to the initial price dispersion. Generally, step-price is a widely used pricing strategy in the used car market. Dealerships that use this strategy sell cars with less mileage, are larger in terms of the number of cars on offer, and are more likely to be an official partner of BMW.

Success of pricing strategies

An important question is how well each individual strategy performs which gives an indication of it fits to the given environment, i.e., its ecological rationality. To compute the performance we use the last price that a dealer advertizes online for a car as an indication for how much the car approximately sells by.[9] From this, the minimum price of the group of matching cars is subtracted which yields the dispersion of the last price. The difference is standardized using the minimum price as a denominator and a mean is computed for each dealer. The results can be seen in table 3 for each of the clusters. The different clusters of dealers vary in the mean number of days that a car is on the lot. For each day, dealers incur variable costs which have to be covered. As an approximation, these are about .01% per day that a car is on the lot of the profit a dealer makes (Löhe, 2010). Multiplying the time until a car sell with .01% yields

[9] An obvious caveat here is that there might be systematic differences between the clusters as to the likelihood that dealers engage in bargaining over the final price with a customer which also can take the form of extras or services provided. This cannot be deduced from the online price.

the costs per cluster. As an auxiliary assumption, we equate profit with the price dispersion of the last price. The dispersion of the last price minus the costs per day on the lot yields the 'earnings' above minimum price per cluster as depicted in table 3.

TABLE 3. 'EARNINGS' PER CAR AND CLUSTER.

	Step-price		**Competitive**		**Wait - price change**		**Wait - no change**	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Dispersion last price	8.9%	5.5%	2.2%	2.4%	6.6%	6.7%	5.1%	8.0%
Time until car sells (days)	49.7	26.3	41.5	26.2	93.2	51.2	32.2	29.6
'Earnings' above min price	8.4%	5.2%	1.8%	2.1%	5.7%	6.2%	4.7%	7.7%

The 'Step-price' cluster is performing best with mean earnings per car above minimum price of 8.4% (SD = 5.2). In contrast, dealers that use a price strategy such that they are part of the 'Competitive' cluster perform worst.

Price points

The interviews indicate that price point is the most relevant theory for the pricing of used car dealers. Price points imply that certain digits, e.g., 9, are more attractive and hence should also be used more frequently in setting a price. If a digit's appearance in a price is random, then each digit is expected to occur with about 10%. However, for the first to the right, 0 accounts for 80.4% of all digits, for the second digit, it accounts for 27.3%. 9 is a prominent second digit with 32.2% and third digit with 42.8% (see table A3 in the appendix).

Note, that the use of price points also implies that market information is ignored to a certain extent. The focus is less on marginally underbidding the competitor but rather on providing the customer with attractive price points.

Price stickiness and dispersion

The literature on sticky prices highlights that it is difficult to identify at which speed a market clearing price would move and therefore to evaluate to what extend prices might indeed be sticky (Blinder et al. 1998). The data on the used car market provides a workaround. Dealers that are part of the 'Competitive'

cluster compete on price aiming to be the cheapest in a group of matching cars. The mean rank their cars achieve is 2.0 (SD = 1.7), with a median of 1.5 where the car with the minimum price is ranked number 1. To illustrate the difference, dealers that are part of the 'Step-price' cluster have a mean rank of 4.7 (SD = 2.4) and a median of 4.4 (see table A2 in the appendix for all clusters and the complete data set). Recall that the median size of a group of matching cars is 5. In order to evaluate the degree of price stickiness the time interval that competitive dealers need to adjust a price is used as approximating the speed at which the market clearing price moves. In the mean, prices of dealers of the 'Competitive' cluster change after 21.1 (SD = 9.0) days, with a median of 20.9 days. In comparison, prices change in the mean at all other dealerships after 41.7 (SD = 25.3) days with a median of 34.6 days (N = 577). Using the median values as a more robust measure, this implies a price stickiness of 65.6% of the prices of dealerships that are not part of the 'Competitive' cluster.

The mean price dispersion of a dealer is calculated using the mean relative difference of the initial price minus the minimum price of the cheapest, matching car. The difference is standardized using the minimum price as a denominator and a mean is computed for each dealer across all prices posted.[10] The mean price dispersion is 7.3% (SD = 6.2), the median is 7.0%.

Discussion

Economic theories have generally assumed that firms use all the available information when setting prices in their pursuit to maximize profits. We show that dealers use readily available market information only very selectively, applying a step-price heuristic: They initially set a high price, which is kept constant for a fixed time interval during which the market is ignored. If the car does not sell during this period, the price is lowered. This procedure is repeated until the car sells. The heuristic is employed by more than half of all used car dealerships as stated in the interviews and found in the online data. 24% of dealerships consult

[10] This procedure deviates from literature that frequently computes the range between the two most extreme price points and neglects other data points (e.g., Brynjolfsson & Smith, 2000), however this would be prone to outliers which are very likely to occur in the used car data set.

the market even only once – when setting the first price of a used car; any subsequent price changes that might become necessary are determined at the first day the car is on offer. Prices are set in such a way that price points feature strongly suggesting that prices 'jump' from one prominent digit to the next. The use of the step-price heuristic is not affected by the degree of competitiveness of the environment dealers operate in; the number of matching cars even positively correlates with initial price dispersion.

The widespread use of the step-price heuristic leads to sticky prices and price dispersion. None of the theories that have been found in previous studies that relate to sticky prices (cost-based pricing, implicit and explicit contracts, co-ordination failure) is rated as relevant in the used car market. The aforementioned price points receives the most approval followed by searching for the best price (Conlisk et al. 1984; Lazear, 1986). This is complemented by results from the regression analysis that dealers who use the step-price strategy tend to be official partner of BMW which entails upholding certain quality standards, to be larger in terms of the number of cars on offer, and to sell cars with less mileage. The first two factors likely serve as a proxy for reputation, diminishing quality uncertainty (Heal, 1976).

A theory that does not receive support is the one by Baye & Morgan (2001) who build on Varian (1980). Price dispersion in the used car market does not stem from randomizing prices, such that consumers cannot learn which one is the most expensive price. Rather those that employ step-price consistently offer the most expensive cars. The higher prices achieved by a step-price strategy can be regarded as a price premium due to a lower quality uncertainty. Being the cheapest in the market is only considered a worthwhile target by the mean dealer after a car has not sold for 180 days. For 27% of interviewees it is not desirable to be the cheapest in the market as they consider these offers as fraudulent. This suggests that a race to the bottom in terms of prices plays only a small role in the used car market. It is rather the search for demand that drives pricing. Also, the importance of price points highlights the focus on demand instead of best responding to the offers of competitors.

The above results highlight that the used car market is not well characterized by standard competitive models. An alternative explanation for the observed pricing behavior is that tacit collusion takes place which allows market participants to reap benefits above market clearing price. Traditionally, collusion is regarded the domain of oligopolies, yet, there has also been evidence that gasoline stations practice collusive pricing, a market characterized by many differentiated firms (Boreinstein & Shepard, 1996). The use of the step-price strategy might thus be the expression of collusive behavior constituting at the same time a well adapted response to the complexity and large degree of uncertainty prevalent. Of those dealers that employ the step-price strategy (N = 423), 61.5 percent are an official partner of BMW which might further facilitate collusion. Yet, in the gasoline market, there are only a handful of prices posted which facilitates to control whether others maintain the collusive equilibrium. Instead, the mean used car dealer has 74 (SD = 45) cars posted online at any one time. In combination with a very large number of dealerships that operate in the market it becomes difficult to detect and punish dealers that deviate from collusive behavior. The current gains from defecting from a collusive equilibrium are likely larger than compared to any future loss due to punishment. Hence, collusive behavior seems unlikely but cannot fully be disregarded as a potential explanation inviting future research.

Price setting in the used car market faces a large degree of complexity and uncertainty. Similar environments can be found for instance in the real estate market (e.g., Levitt & Syverson, 2008), or with 'one-of-a-kind' products such as designer gowns (Lazear, 1986). Demand uncertainty can also be an issue with homogenous products, particularly at their introduction which has been shown even in the case of pharmaceuticals that undergo extensive testing and market research before their introduction. Many drugs differ in their idiosyncratic side effects that might show up only when used in a wider population. Even after extensive research the firm does not know the value of its product exactly and needs to search for the best price (Crawford & Shum, 2005).

How efficient actually employed pricing strategies indeed are has recently been questioned by Fabiani et al. (2007). They find that two thirds of firms employ what they classify as a suboptimal pricing rule. However, the present study shows

that a fast and frugal pricing heuristic that might by conventional standards of as-if models be considered to be a suboptimal strategy in fact performs very well.

How does one best approach such environments as the used car market from a modeling perspective? The management literature suggests that the problem in its entirety is mathematically intractable. However, the documented step price strategy provides an avenue that cuts through this and can serves as a base for model development. This can be combined with approaches provided by Lazear (1986) who models price discovery in the face of an uncertain demand and Conlisk, Gerstner, & Sobel (1984) who address the issue of strategic customers. At the same time this raises the question in how far a model that uses a maximization approach can predict behavior well in the given environment (Gigerenzer & Brighton, 2009). An alternative is the approach by Selten (1998) and the use of the aspiration adaptation theory that models a search process.

Competitive model testing is likely to provide further insights with regards to the issue of the predictive power of models that employ maximization. Aviv & Pazgal (2005) and Mersereau & Zhang (2011) are able to address the problem of pricing by reducing the information load their models deal with. Yet, these are still computationally highly intensive and rely on a maximization process. Both strategies could be tested against the non-maximizing step-price strategy in terms of their performance. Such a model testing competition would also give further insights into the degree of price stickiness and dispersion in a given market or even at the very same dealership.

Bibliography

Akerlof, G. A., & Yellen, J. L. (1985). A Near-Rational Model of the Business Cycle, With Wage and Price Inertia. *The Quarterly Journal of Economics*, *100*, 823–838.

Allen, F. (1988). A Theory of Price Rigidities When Quality is Unobservable. *The Review of Economic Studies*, *55*(1), 139–151.

Amirault, D., Kwan, C. C., Wilkinson, G., & Canada, B. du. (2006). *Survey of price-setting behaviour of canadian companies*. Bank of Canada.

Apel, M., Friberg, R., & Hallsten, K. (2005). Microfoundations of Macroeconomic Price Adjustment: Survey Evidence from Swedish Firms. *Journal of Money, Credit and Banking*, *37*(2), 313–338.

Aviv, Y., & Pazgal, A. (2005). A Partially Observed Markov Decision Process for Dynamic Pricing. *Management Science*, *51*(9), 1400 –1416.

Balvers, R. J., & Cosimano, T. F. (1990). Actively Learning About Demand and the Dynamics of Price Adjustment. *The Economic Journal*, *100*(402), 882–898.

Baye, M. R., & Morgan, J. (2001). Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets. *The American Economic Review*, *91*(3), 454–474.

Baye, M. R., Morgan, J., & Scholten, P. (2006). Information, search, and price dispersion. In T. Hendershott (Ed.), *Handbook on economics and information systems* (pp. 323–376). Elsevier.

Bils, M. (1987). The Cyclical Behavior of Marginal Cost and Price. *The American Economic Review*, *77*(5), 838–855.

Binmore, K. (2007). Rational Decisions in Large Worlds. *Annales d'Économie et de Statistique*, (86), 25–41.

Blinder, A. S. (1991). Why are Prices Sticky? Preliminary Results from an Interview Study. *The American Economic Review*, *81*(2), 89–96.

Blinder, A. S., Canetti, E. R. D., Lebow, D. E., & Rudd, J. B. (1998). *Asking about prices - A new approach to understanding price stickiness*. New York: Russell Sage Foundation.

BMW Group. (2011). BMW Group Annual Report 2010.

Boreinstein, S., & Shepard, A. (1996). Dynamic Pricing in Retail Gasoline Markets. *RAND Journal of Economics*, *27*(3), 429–451.

Brynjolfsson, E., & Smith, M. D. (2000). Frictionless Commerce? A Comparison of Internet and Conventional Retailers. *Management Science*, *46*(4), 563–585.

Bulow, J. I. (1982). Durable-Goods Monopolists. *Journal of Political Economy*, *90*(2), 314–332.

Carlton, D. W. (1986). The Rigidity of Prices. *The American Economic Review*, *76*(4), 637–658.

Coase, R. H. (1972). Durability and Monopoly. *Journal of Law and Economics*, *15*(1), 143–149.

Conlisk, J., Gerstner, E., & Sobel, J. (1984). Cyclic Pricing by a Durable Goods Monopolist. *The Quarterly Journal of Economics*, *99*(3), 489–505.

Cooper, R., & John, A. (1988). Coordinating Coordination Failures in Keynesian Models. *The Quarterly Journal of Economics*, *103*(3), 441–463.

Crawford, G. S., & Shum, M. (2005). Uncertainty and Learning in Pharmaceutical Demand. *Econometrica*, *73*(4), 1137–1173.

Dudenhöffer, F., & Schadowski, A. (2011). Die Online-Gebrauchtwagen-Börsen in Europa. Universität Duisburg-Essen, CAR-Center Automotive Research.

Ellison, G. (2006). Bounded Rationality in Industrial Organization. *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress* (Vol. 2, p. 142–74).

Fabiani, S., Loupias, C. S., Martins, F., & Sabbatini, R. (2007). *Pricing Decisions in the Euro Area: How Firms Set Prices and Why*. Oxford University Press.

Fischer, S. (1977). Long-Term Contracts, Rational Expectations, and the Optimal Money Supply Rule. *Journal of Political Economy*, *85*(1), 191–205.

Friedman, M. (1953). *Essays in positive economics*. University of Chicago Press.

Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why Biased Minds Make Better Inferences. *Topics in Cognitive Science*, *1*(1), 107–143. doi:10.1111/j.1756-8765.2008.01006.x

Gigerenzer, G., Todd, P. M., & the ABC Research Group. (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press.

Gode, D. K., & Sunder, S. (1993). Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality. *The Journal of Political Economy*, *101*(1), 119–137.

Hall, S., Walsh, M., & Yates, A. (2000). Are UK companies' prices sticky? *Oxford Economic Papers*, *52*(3), 425 –446.

Heal, G. (1976). Do Bad Products Drive Out Good? *The Quarterly Journal of Economics*, *90*(3), 499–502.

Kelman, M. (2011). *The heuristics debate*. Oxford University Press.

Lach, S. (2002). Existence and Persistence of Price Dispersion: An Empirical Analysis. *Review of Economics and Statistics*, *84*(3), 433–444.

Lazear, E. P. (1986). Retail Pricing and Clearance Sales. *The American Economic Review*, *76*(1), 14–32.

Levitt, S. D., & Syverson, C. (2008). Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions. *Review of Economics and Statistics*, *90*(4), 599–611. doi:10.1162/rest.90.4.599

Levy, D., Lee, D., Chen, H. (Allan), Kauffman, R. J., & Bergen, M. (2011). Price Points and Price Rigidity. *Review of Economics and Statistics*, *93*(4), 1417–1431.

Löhe, J. (2010). Studie zur Professionalisierung im Gebrauchtwagengeschäft: Handlungsfelder und Potentiale. Working paper, European Business School,.

Lovejoy, W. S. (1991). Computationally Feasible Bounds for Partially Observed Markov Decision Processes. *Operations Research*, *39*(1), 162–175.

Maccini, L. J. (1973). On optimal delivery lags. *Journal of Economic Theory*, *6*(2), 107–125.

Mankiw, N. G. (1985). Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly. *The Quarterly Journal of Economics*, *100*(2), 529–537.

Mersereau, A. J., & Zhang, D. (2011). *Markdown pricing with an unknown fraction of strategic consumers*. Working paper.

Michel, W. (2009). DAT Report 2009.

Okun, A. M. (1981). *Prices and Quantities: A Macroeconomic Analysis*. Brookings.

Rothschild, M. (1974). A two-armed bandit theory of market pricing. *Journal of Economic Theory*, *9*(2), 185–202. doi:10.1016/0022-0531(74)90066-0

Savage, L. J. (1954). *The foundations of statistics*. Courier Dover Publications.

Selten, R. (1998). Aspiration Adaptation Theory. *Journal of Mathematical Psychology*, *42*(2–3), 191–214. doi:10.1006/jmps.1997.1205

Simon, H. A. (1962). New Developments in the Theory of the Firm. *The American Economic Review*, *52*(2), 1–15.

Smith, V. L. (2008). *Rationality in Economics: Constructivist and Ecological Forms* (1st ed.). Cambridge University Press.

Tauchen, G. E., & Pitts, M. (1983). The Price Variability-Volume Relationship on Speculative Markets. *Econometrica*, *51*(2), 485–505.

Telser, L. G., & Higinbotham, H. N. (1977). Organized Futures Markets: Costs and Benefits. *Journal of Political Economy*, *85*(5), 969–1000.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124 –1131.

Varian, H. R. (1980). A Model of Sales. *The American Economic Review*, *70*(4), 651–659.

Appendix

A1. Interview car dealers

- *Text in italics and underlined is only meant for the interviewer*
- *The interview is recorded on audio tape and transcribed*
- *The interview lasts about 20 minutes*
- *Each interview is given an ID. This ID is identical to the ID from the online data set in order to compare statements made in the interview to the actual pricing behavior.*
- *UC = used car*

The Max Planck Institute for Human Development is interviewing second hand car dealers for the Germany-wide study "Pricing behavior of second hand car dealers". You will be asked to respond to questions regarding how you set prices for UCs in order to contribute to the success of your enterprise. The aim of the study is to document how UC dealers price. Anything you say is anonymized and analyzed only on an aggregate level. For this the audio recordings will be transcribed and destroyed thereafter. The final analysis is based on the interviews and data from the online market Autoscout.de

I. General Questions

1. What is your function in the dealership: _____
2. a) How many cars, old and new, are currently on offer in this dealership: _____

   b) How many of these are UCs: _____

   c) How many UCs are BMWs: _____
3. How large is the degree of returning customers, i.e., those that have bought at least their last car at the dealership: _____
4. a) Is the dealership part of a larger organization and if so, how many branches does this have: _____

   b) *if a. > 1*: Are the prices for UCs

   ☐ centrally done ☐ individually done at each dealership
5. Since when do you price UCs: _____

II. Prices – specific questions

6. Initial price: a) What do you do if you have to price a UC the first time:

   _____

   _____

   b) Which information is important to determine the initial price:

   _____

   _____

7. You are faced with the situation where your competitors offer a similar UC. The initial price of your UC is

   ☐ i) lower than the price of the competitor

   ☐ ii) the same as the competitor

   ☐ iii) higher than the competitor

8. What is the geographical size of the market you are selling your product (*km in diameter*) _____

9. a) Do you commonly change prices: *yes or no*

   b) *if a = yes:* How do you proceed if you have to change a price?

   _____

   _____

   c) When do you change a price: *(time interval, date, event)*:_____

   d) By how much do you change the price: *(in % or €)*:_____

   e) If you change the price of a UC, do you collect information on the current price of a similar car of a competitor each time you do this:

   _____

   f) Is there a point of time when you offer the UC cheaper than similar cars of your competitors, hence your car is ranked first in the online market:

   _____

   g) *if f = no:* Why not:

   _____

10. a) Are you using the same pricing strategy for all (most) of your UCs: *yes or no*

    b1) *if a = no:* Which criterion do you apply to differentiate between UCs, which pricing strategy do you apply and how often do you apply them?

    _____

11. a) Are all UCs on the lot also advertized on the internet market? *yes or no*
    b) *if b = no:* Why not: _____
12. How large is the share of customers that ***only*** focus on price for their buying decision***:***_____

III. **Economic Theories** - How important is any of the following theories with regards to the pricing of second hand cars in your daily business.

1 = totally unimportant 2 = of minor importance
3 = moderately important 4 = very important

I. (*cost based pricing*) A different idea holds that prices depend mainly on the costs of labor and of materials and supplies that companies buy from other companies. Firms are thought to delay price increases until their costs rise, which may take a while. But then they raise selling prices promptly (as in Blinder et al. (1998) question B6).

II. (*implicit contracts*) Another idea has been suggested for cases in which price increases are not prohibited by explicit contracts. The idea is that firms have implicit understandings with their customers – who expect the firms not to take advantage of the situation by raising prices when the market is tight (as in Blinder et al. (1998) question B2).

III. (*explicit contracts*) One idea is that many goods are sold under explicit contractual agreements that set prices in advance, so firms are not free to raise prices while contrasts remain in force (as in Blinder et al. (1998) question B1).

IV. *(co-ordination failure)* The next idea is that firms would often like to change their prices, but are afraid to get out of line with what they expect competitors to charge. They do not want to be the first ones to raise prices. But, when competing goods rise in price, firms raise their own prices promptly (as in Blinder et al. (1998) question B10).

V. (*non-price competition*) The idea here is that firms don't cut prices much when demand falls because price is just one of several elements that matter to buyers. More frequently, they shorten delivery lags, make greater selling

efforts, improve service, or improve product quality (as in Blinder et al. (1998) question B12).

VI. (*price points*) Another idea is that particular threshold prices are more attractive to customers than other prices. For example a dealership might think that a used car sells much better at 9.990€ than 10.050€ (as in Blinder et al. (1998) question B4).

VII. (*judging quality by price*) One idea is that firms hesitate to reduce their prices because they fear that customers will interpret a price cut as a signal that the quality of the product has been reduced (as in Blinder et al. (1998) question B3).

VIII. *(menu costs)* Another idea is that the act of changing prices entails special costs in itself, so firms hesitate to change prices too frequently or buy too much. The costs we have in mind are not production costs but costs like printing a new catalogue, price lists, etc. or hidden costs like loss of future sales by antagonizing customers, decision making time of executives, problems with sales person and so on (as in Blinder et al. (1998) question B8).

IX. (*search*) A theory says that customers know approximately the price range that customers are willing to pay. In order to achieve the best possible price, firms start with a high price and reduce it regularly. Sale occurs once a customer is willing to pay the advertised price.

A2. Further statistics

TABLE A1. RANKING OF THEORIES BY INTERVIEWEES WHO USE STEP-PRICE (N = 38) AND WHO DO NOT (N = 17).

	Step-price			Not step-price		
	Mean	Median	SD	Mean	Median	SD
Price points	3.4	4.0	1.0	2.7	3.0	1.2
Search	3.0	3.0	1.0	2.5	2.0	1.1
Non-price competition	1.8	2.0	0.6	1.8	2.0	0.6
Cost-based pricing	1.9	1.0	1.4	1.9	1.5	1.1
Implicit contracts	1.8	1.5	0.9	1.7	1.0	0.9
Co-ordination failure	1.6	1.0	0.8	1.7	1.0	0.9
Menu costs	1.2	1.0	0.8	1.6	1.0	0.9
Judging quality by price	1.4	1.0	0.5	1.4	1.0	0.6
Explicit contracts	1.1	1.0	0.4	1.8	1.0	1.2

TABLE A2. INDEPENDENT VARIABLES: MEAN VALUE AND STANDARD DEVIATION FOR EACH CLUSTER AND THE COMPLETE DATA SET (N = 748).

	Complete data			Step-price			Competitive			Wait - price change			Wait - no change		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Further price characteristics															
Number of price changes per car	.8	.6	.8	.9	.7	.8	1.0	.8	.9	.6	.5	.4	.0	.0	.0
Rank (1=cheapest)	3.6	3.1	2.5	4.6	4.4	2.4	2.0	1.5	1.7	3.1	2.6	2.3	2.2	1.4	2.1
Price (€)	23,482	24,341	8,954	26,470	27,126	6,641	20,074	18,472	8,729	21,153	19,810	9,586	18,247	14,994	11,841
Characteristics of dealerships															
Number of 320 and 730	32.7	13.0	65.5	48.5	27.0	80.4	16.3	7.0	24.2	18.7	6.0	52.5	3.2	2.0	3.9
Proportion of 320	93.7%	100.0%	14.1%	93.7%	96.7%	8.7%	94.2%	100.0%	15.7%	93.4%	100.0%	16.9%	92.9%	100.0%	23.7%
Cars online at the same time	74.0	45.0	117.9	93.4	60.0	141.2	50.3	30.0	55.8	80.1	53.0	143.3	34.4	25.0	34.6
Competitive setting - number of cars	800	749	473	808	749	476	790	758	419	813	725	571	785	749	503
Mileage (km)	69,102	57,788	37,914	57,859	49,991	28,450	81,136	76,577	40,023	80,657	76,507	45,132	88,931	85,690	47,362
Share of cars with extra warranty*	23.9%	0.0%	34.4%	31.1%	11.3%	35.2%	15.9%	0.0%	31.4%	15.8%	0.0%	30.7%	11.9%	0.0%	30.5%
Part of network*	18.6%	0.0%	38.9%	24.8%	0.0%	43.2%	11.8%	0.0%	32.3%	11.6%	0.0%	32.4%	8.1%	0.0%	27.4%
Official partner of BMW*	47.1%	0.0%	49.9%	61.5%	100.0%	48.7%	28.1%	0.0%	45.1%	34.9%	0.0%	48.2%	26.1%	0.0%	44.1%

* This is not a mean value but refers to the proportion of dealers of the total data / per cluster.

TABLE A3. FREQUENCY OF PRICE DIGITS FROM STARTING FROM THE RIGHT END OF THE PRICE.

Digits - from right	1st	2nd	3rd	4th	5th
0	80.4%	27.3%	1.4%	7.6%	0.0%
1	2.2%	2.0%	1.1%	9.2%	27.1%
2	1.9%	1.1%	3.0%	9.8%	32.8%
3	0.9%	3.3%	3.7%	9.2%	24.6%
4	1.1%	3.7%	13.0%	10.0%	5.7%
5	4.0%	13.6%	7.0%	8.9%	3.0%
6	0.8%	2.5%	4.4%	9.9%	3.0%
7	1.7%	3.5%	7.0%	10.6%	1.6%
8	1.6%	10.8%	16.5%	10.6%	1.2%
9	5.5%	32.2%	42.8%	14.2%	1.0%