Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Franziska Tausch

The Distributional Impact of Risk Heterogeneity, Risk Responsibility and Control

PhD Thesis 2014-002

The distributional impact of risk heterogeneity, risk responsibility and control

This book was typeset by the author using LaTeX.

Published by Datawyse | Universitaire Pers Maastricht
ISBN 978 94 6159 324 5

The distributional impact of risk heterogeneity, risk responsibility and control

Dissertation

to obtain the degree of Doctor at
Maastricht University,
on the authority of the Rector Magnificus,
Prof. dr. L.L.G. Soete,
in accordance with the decision of the Board of Deans,
to be defended in public
on Friday, April 11, 2014, at 10:00 hours

by

Franziska Tausch

UNIVERSITAIRE PERS MAASTRICHT

Supervisors:
Prof. dr. Arno Riedl
Prof. dr. Jan Potters (Tilburg University)

Assessment Committee:
Prof. dr. Thomas Dohmen (chairman)
Dr. Matthew Embrey
Prof. dr. Jean-Robert Tyran (University of Vienna)

Financial support by the Network for Studies on Pensions, Aging and Retirement (Netspar) is gratefully acknowledged.

Contents

Chapter 1

Introduction

In contrast to classical economics, behavioral economics accounts for individuals' bounded rationality as well as for insights from psychology when predicting economic behavior. Plenty of empirical evidence suggests that individuals' decisions are not solely inspired by self interest as suggested by the classical concept of homo economicus. Instead, many individuals exhibit social, or other regarding preferences. They do not only care about their own material payoffs from social and economic interactions, but they also care about the payoffs of their interaction partners.

In **Chapter 2** of my thesis (joint with Arno Riedl and Jan Potters) I review experimental research that deals with individuals' preferences for (re)distribution. We distinguish between three types of situations. The first deals with distributional preferences *behind* a veil of ignorance. Decision makers do not know their own income position and are thus largely impartial and unaffected by their immediate self-interest. We discuss experiments and vignette studies that are informative with respect to normative justice principles in connection with e.g. earned rights and risk taking. In the second type of situation, individuals make choices *in front* of the veil of ignorance and know their income position. Here we discuss the role of stake sizes, the income generating process and decision power, social identity and social distance, gender, age, ideology as well as the role of institutions. Finally, we consider situations in which income is determined by interdependent rather than individual choices. Individuals make decisions in strategic settings in which they are mutually dependent. Experiments on risk sharing and insurance as well as on intergenerational transfers are discussed. In the closing parts of the paper we relate the experimental results to the redistribution issues of insurance and pensions. While the reviewed evidence clearly shows that people share some basic willingness to support redistribution in general, it also points to the limits of this support. We conclude that the structure and distributional consequences of solidarity-based pension schemes have to be in line with generally shared fairness norms, as otherwise, the schemes will lose societal support and open the door to a host of adverse consequences.

The essential mechanism of any kind of insurance system like pension or health insurance is the sharing of risks. It diminishes individuals' vulnerability to probabilistic events that negatively affect their financial situation.

In the pension context individuals face the risk of outliving their pension savings; in the health context they are exposed to the risk of getting ill which results in the need for costly medication. Risk sharing implies that money is redistributed from the 'lucky' individuals for whom no undesirable event occurs to those that are less lucky. The acceptance of the resulting redistributive transfers is crucial for the support and sustainability of risk sharing arrangements. The more risk is shared the more inequalities in incomes are reduced and the stronger the financial safety net that is provided.

In **Chapter 3** of my thesis (joint with Arno Riedl and Jan Potters) I investigate whether adverse selection hampers the effectiveness of voluntary risk sharing and how differences in risk profiles affect adverse selection. To this end, we experimentally investigate individuals' willingness to share risks with others. Across treatments we vary how risk profiles differ between individuals. Four treatments are implemented that differ with respect to stochastic dominance among individuals' risks. In the benchmark treatment risks are identical whereas in the other treatments risk heterogeneity is implemented. All treatments can be ranked with respect to the likelihood of adverse selection occurring. We find strong evidence for adverse selection if individuals' risk profiles can be ranked according to first-order stochastic dominance and only little evidence for adverse selection if risk profiles can only be ranked according to mean-preserving spreads. We observe the same pattern also for anticipated adverse selection. These results suggest that the degree to which adverse selection erodes voluntary risk sharing arrangements, and thus redistribution between individuals, crucially depends on the form of risk heterogeneity.

The extent to which individuals are willing to engage in risk sharing may not only depend on individuals' relative risk exposure, but also on how the risk exposure came about.

In **Chapter 4** of my thesis (joint with Elena Cettolin) I investigate the role that responsibility for risk exposure plays in a risk sharing context. We hypothesize that when people are responsible for the type of risk they face, individuals' willingness to share risk is low due to dampened redistribution motives, and investigate this conjecture with a laboratory experiment. Responsibility is created by allowing participants to choose between two different risky lotteries before they decide how much risk they share with a

randomly matched partner. Risk sharing is then compared to a treatment where risk exposure is randomly assigned. We find that average risk sharing does not depend on whether individuals can control their risk exposure. However, we observe that when individuals are responsible for their risk exposure, risk sharing decisions are systematically conditioned on the risk exposure of the sharing partner, whereas this is not the case when risk exposure is random. In particular, when risk exposure is deliberate individuals are less willing to share risk with high risk takers as compared to low risk takers.

Another context in which distributional preferences may interact with preceding decisions of the interaction partner is that of principal agent relationships. The principals' outcomes typically depend on the agents' effort provision. Providing effort, however, comes at a cost for the agent such that his payoff maximization is not in line with the principal's interest. The principal thus has an incentive to control the agent in order to enforce a minimum effort provision.

In **Chapter 5** I investigate how agents' effort provision is affected by the imposition of control. To this end, I implement a principal agent game with two principals. Before the agent decides how much effort to provide (how many points to distribute) to each principal, a 'controller' can decide to impose a minimum effort requirement for only *one* of the principals. I compare agent's behavioral reaction to two types of control. The first type is internal, implying that the principal for whom control takes effect is the controller. The second type is external, implying that an individual whose earnings are unrelated to the agent's effort makes the control decision. When the principal himself is the controller I find that a considerable fraction of agents reduce their effort provision towards him when controlled as compared to when there is no control. The unconstrained effort provided to the second principal, however, is not systematically affected by the control decision. The average amount of effort provided does not depend on the controller's control decision. In contrast to the case of internal control, the fraction of agents reacting negatively to control is negligible when the controller is external. In that case effort provision is significantly higher towards both principals when there is control as compared to when there is no control. The results show that control only has detrimental effects if the controller is the princi-

pal who profits (suffers) from high (low) effort provision of the agent. This suggests that in the aggregate, control can be effective in enforcing high effort provision, if an external third party, like e.g. the government or a consultancy, makes the control decision.

Chapter 2

Preferences for Redistribution and Pensions. What Can We Learn from Experiments?*

*This chapter is joint work with Jan Potters and Arno Riedl and has been published in the *Journal of Pension Economics and Finance* (doi:10.1017/S1474747212000388).

2.1 Introduction

'A collective pension is always solidarity', according to an advertisement of the world's biggest pension fund (APB). Obviously, the pension fund believes that people value solidarity positively. Broadly speaking, solidarity refers to 'a positive sense of shared fate between individuals or groups. That is, a situation where social relationships center on the stronger helping the weaker or on promoting the communal interest' (Van der Lecq and Steenbeek, 2007, p. 4). In the domain of pensions, solidarity can take place at different levels. A distinction can be made between risk solidarity, subsidizing solidarity and income solidarity. *Risk solidarity* is a consequence of risk sharing, and it implies that *ex post* the lucky support the unlucky. *Subsidizing solidarity* involves *ex-ante* value transfers from one group to another - as is the case, for example, when longevity risk is expected to be larger for one group (women) than for another (men). *Income solidarity* usually implies that income is redistributed from the rich to the poor - as is the case, for instance, for old-age social security (AOW) in the Netherlands where contributions are income-dependent, while benefits are not. Whatever its form, however, solidarity is always about redistribution (Centraal Planbureau, 2000).

An important question is whether and why people support the redistribution embodied in collective pension schemes. Some forms of support may be rooted in self-interest, such as redistribution resulting from risk solidarity, which is mutually advantageous when people are risk averse. Redistribution due to subsidizing- and income solidarity is advantageous for those on the receiving end. Hence, self-interest can explain these forms of redistribution, if one assumes that the groups who receive have the political power to pursue their interests at the expense of those who pay. Apart from the fact that this is a tenuous assumption, indeed, an attempt to explain redistribution merely on the basis of self-interest is too restricted a perspective. One should not rule out the possibility that many people do in fact have *social preferences* - that is, a genuine concern for the welfare of others and a preference for a just and fair distribution of incomes and risks. Increasing numbers of economists (or 'even economists', one could say) believe this to be the case. This belief is at least partly based on experimental evidence that has been collected in the last two decades or so.

This paper reviews the experimental literature on social preferences, and discusses the implications for redistribution and pensions. We should mention, however, that few experimental studies directly address solidarity with regard to pension schemes. For example, several studies deal with plain distribution and redistribution, but few of them focus on such issues as subsidizing solidarity in risk sharing or solidarity across the generations. Still, we believe that the results from this literature can add some empirical evidence to pension reform discussions, which are all too often based on mere speculation about what people really prefer. Moreover, preferences regarding redistribution are important not only for debates about pensions but also for fiscal policy and the welfare state, including health care, unemployment insurance, disability insurance and poverty alleviation.

The remainder of the paper is organized as follows. Section 2.2 provides a brief introduction to the methodology of experimental economics. Section 2.3 is the main body of the paper, providing an extensive review of the experimental literature that deals with social preferences and redistribution. Section 2.4 gives a summary and our interpretation of the main results. Section 2.5 outlines the important missing elements in the experimental literature with respect to issues of pension solidarity. Finally, Section 2.6 concludes and presents the main implications.

2.2 The method of experimental economics

In an economic experiment, human subjects make decisions in a controlled environment. The typical procedure is that participants (usually students) are invited to an experimental laboratory where they receive instructions that provide details about the rules in the experiment. Specifically, they explain how one's earnings will be affected by own decisions and, possibly, by decisions of other subjects and chance. To ensure proper understanding of the rules and incentives, comprehension questions are asked and often a practice round is run. At the end of the experiment subjects are paid out their earnings confidentially in cash.

The key issue of any experiment is control. That is the experimenter has full control over the decision environment and can change environmental variables in a controlled way that suits the research question. Experiments

are used for a variety of research purposes. One prominent use is testing the predictions of economic models which are by their nature an abstraction of the complexities of reality. In an experiment, in contrast to a field setting, the theoretical model and the actual decision environment can be brought close together. The data of the experiment can thus provide a clean test of the economic model. This is what is sometimes called 'testing a theory on its own domain'. A related advantage is that one variable can be changed at a time which is particularly important if one wants to make causal inferences. Experiments can also be used for 'testbedding'. Just as scale models of airplanes are tested in a wind tunnel, one can implement different policies and institutions in a controlled setting and compare their performance. Experiments have been used, for example, to evaluate different tax systems (see, e.g., Riedl and Van Winden, 2007, Riedl and van Winden, 2012) and various auction designs (for a recent overview of policy-related experiments, see Normann and Ricciuti, 2009). Another important reason for using experiments is that they make it possible to explore and measure behavioural parameters such as risk attitudes, discount rates, probability weighting, or predictive abilities in an incentive-compatible way.

This paper is concerned with social preferences, which are broadly defined as the manner and degree to which people care about the well-being of others and about the aggregate outcome. An important feature of economic experiments is that participants can earn money, and that the money they earn depends on their decisions (which is - next to the 'no deception' rule - one of key differences with most experiments in psychology). This ensures that subjects are motivated to think about their decisions carefully and to make decisions that reflect their true preferences. This is particularly important for studies of social preferences and pro-social behaviour, because in surveys people may be tempted to give socially desirable answers. In an experiment, however, such social responses have material consequences. In other words, participants are forced to put their money where their mouth is.

An important question is whether experimental results can be generalized. There are two issues related to this question: The concern that laboratory experiments are too simple relative to the environment of interest in the outside world (environmental validity) and the concern that the chosen

subjects are not representative (population validity). With regard to the first concern, it is important to realize that the main purpose of an experiment - just like in a theoretical model - is to identify the essential environmental variables for the research question at hand. General theoretical principles (self-interest, rationality, maximization and equilibrium) can be and often are tested with rather abstract experimental designs, whereas in the case of test-bed experiments, more effort is made to minimize the distance between the experimental design and the specific environment of interest. Moreover, experiments are ideally suited to gradually increase the complexity of the environment (principle of decreasing abstraction). This makes it possible to trace precisely which factor is responsible for a particular change in the observed outcomes.

The second issue of external validity is the choice of experimental subjects. University students are often used as subjects because they are easily available and have relatively low opportunity costs. But the question is whether their behaviour is indicative of that of 'real people'. To investigate this question, researchers have carried out a number of selective replications of experiments using the relevant subjects as participants (the general population, voters, employees and managers, for example). Even though some differences are found, the results of these studies indicate that the general patterns of behavior of 'real people' usually correspond remarkably well with those found with student subjects (Fréchette, 2011). Having said that, it must surely be acknowledged that the experimental method, like any method, has its limitations. Experiments are no panacea, but a valuable supplementary source of information. Generally, one can say that experimental results are most convincing when they are accompanied by theoretical insights and observations from the field.

2.3 Experiments on income distribution and redistribution

In the experimental literature on 'other-regarding' preferences, three different kinds of settings (designs) can be distinguished. Inspired by Harsanyi (1955) and Rawls (1971), the first setting asks individuals to make decisions

behind a veil of ignorance: this prevents them from knowing their own income position or even their own abilities. The goal is to assess the principles of distributive justice that people uphold when they are largely impartial to the outcome and not affected by their immediate self-interest. These experiments are discussed in Section 2.3.1. In the second setting, individuals make choices *in front of the veil of ignorance*: thus, they know whether they occupy a relatively advantaged or disadvantaged position. As a consequence, distributional preferences will be affected by self-interest. As we see, however, for many people self-interest is not the only guide for their decisions. These experiments are discussed in Section 2.3.2. Finally, in Section 2.3.3 we review experiments in which individuals make decisions in strategic settings. The key feature here is that individuals interact with each other and are mutually dependent. An important question is whether people are willing and able to cooperate when there is tension between individual interest and collective interest.

2.3.1 Preferences regarding income distributions: behind a veil of ignorance

This section reviews experiments investigating the principles of distributive justice to which people adhere. What preferences do individuals have concerning income distributions when they are not biased by self-interest? This matters, because policies that are aligned with generally shared principles are likely to be accepted more easily than those that are opposed to them. Principles of justice are hard to assess in the field, for the simple reason that every individual knows his or her position in society (age, gender, skill and social background). In particular, people know their position in the income distribution, and can by and large predict their absolute and relative future income, including the risks they face. This means that notions of justice that are expressed by people will unavoidably be coloured by self-interest. However, 'objective' justice principles should relate to a situation in which people do not (yet) know their actual position, or, alternatively, a situation in which they are impartial to the outcome.

Experiments investigating principles of justice

One of the first such experimental studies was conducted by Frohlich et al. (1987), followed up by Frohlich and Oppenheimer (1990). Their experiments were carried out with the aim to implement the original position (i.e., behind the veil of ignorance) in the laboratory. In these experiments, students formed small societies in which they had to make *ex-ante* decisions about the different distributive rules to be implemented in the society that they were going to be part of, without knowing what their *ex-post* absolute and relative income position in this society would be. Specifically, in the experiment of Frohlich et al. (1987), the participants had to discuss and unanimously choose one of four distributive principles that would be actually implemented after the determination and announcement of each subject's income position in the society. The four investigated distributive principles were as follows: the principle of maximizing the well-being of the worst-off (Rawls, 1971), the principle of maximizing average well-being (utilitarian), and two constrained forms of maximizing average well-being. When deciding on the distributive principle, participants knew that afterwards they would be randomly allocated to an income class and would earn an amount that depended on that income class and the chosen distributive principle. The main result was that, as a rule, virtually all participants chose a principle that maximized average income with some lower bound on the minimum income that the (*ex-post*) worst-off participant would receive. Hence, there was a preference for a utilitarian society with some safety net, where the choice of a safety net could be ascribed to risk aversion of the participants. In the follow-up study, Frohlich and Oppenheimer extended that set-up to economies with production, and found qualitatively similar results.

Herne and Suojanen (2004) investigate the behaviour of participants for two different original positions: first, the Rawlsian original position behind the veil of ignorance, and second, the Scanlonian original position, which consists of negotiating parties that have full knowledge of their personal characteristics as well as economic and social circumstances, equal bargaining power, and a desire to reach agreement that no one could reasonably reject. Interestingly, the authors found that the Rawlsian outcome was implemented much more often when there was no veil of ignorance (60%) than

when there was (14%). In line with earlier results, however, the most popular distributive principle (62%) behind a veil of ignorance was a utilitarian allocation with a constraint guaranteeing some minimum income for the worst-off.

Distribution choices by a benevolent dictator - vignette studies

Besides implementing the original position, a different way to generate impartiality is to ask participants to make choices that affect others but not themselves. Hence, participants make choices as an impartial referee or - as it is sometimes called - a benevolent dictator. An early example is the study of Yaari and Bar-Hillel (1984). Student respondents are confronted with different scenarios of how to distribute a bundle of commodities in a simple exchange economy. Such surveys are sometimes called *vignette studies*, and often do not involve monetary stakes (and in this sense are not economic experiments). Still, they can generate valuable insights. One of the main interests in this study was under what circumstances a departure from the equal division will occur, which is a very natural and widely accepted justice norm in situations where the engaged agents are symmetric in all relevant aspects. The authors argue that a departure from equal division requires a justification. Accordingly, the investigated scenarios are asymmetric with respect to needs or tastes. Subjects are asked how they would allocate 12 grapefruits (x_1) and 12 avocados (x_2) over Jones and Smith, when Jones' utility function is $u_J = 100x_1$ while that of Smith is $u_S = 20x_1 + 20x_2$. In the scenario in which the utility functions describe the nutritional needs of the individuals, the majority of the subjects prefer the allocation (4, 0) for Jones and (8, 12) for Smith, yielding equal utilities. However, in the scenario in which the utility functions reflect tastes (liking and disliking), the answers are mostly in favour of (12, 0) for Jones and (0, 12) for Smith.

A main finding of this research is that differences in needs weigh much heavier than differences in tastes do as an argument to depart from the equal division. Specifically, in cases of asymmetry in needs, the Rawlsian criterion of maximizing the well-being of the worst-off is chosen most often, whereas in cases of asymmetry in tastes the utilitarian principle of maximization of the sum (or average) of individual utilities is the most popular

choice of the uninvolved student respondents.

Subsequent research using the vignette technique for eliciting principles of distributive justice has introduced *production* into the environment. Schokkaert and Overlaet (1989) compare two scenarios: one in which production depends on effort, and one in which production depends on abilities. They find that 'differences [in effort] completely overrule all other reasons for income differences' (p. 31). Effort differences are seen as morally more just arguments for income differences than are differences in innate abilities. Schokkaert and Capeau (1991) replicate this finding with respondents from the Flemish working population.

Konow (1996) takes up these results and formulates a theory of fairness, which tries to characterize the fairness values people share and to isolate these values from situation-specific contexts. Specifically, the author proposes what he calls the *Accountability Principle* as a general rule of fairness. This principle basically says that a person's fair share should vary with the variables he or she can control (e.g. work effort) - but not with variables that he or she cannot control (e.g. genetic differences). Konow (1996) validates his theory with telephone interviews and written responses to hypothetical scenarios that systematically vary controllable and non-controllable variables. Faravelli (2007) investigates whether support for certain principles of distributive justice (egalitarianism, Rawlsian maximin, utilitarianism and utilitarianism with a floor constraint) varies with the responsibility that individuals bear for the produced outcome. One context was neutral ; in a second context, one individual produced less because of a physical handicap; in a third context, the individual produced less because of little effort. The fairness judgements clearly varied with the context. The less productive individual is relatively favoured (i.e., the maximin principle is chosen) if he or she has a handicap, but is relatively disfavoured (i.e., the utilitarian principle is chosen) if he or she is lazy.

Distributional choices by a benevolent dictator - experimental economic studies

Undoubtedly, important insights can be gained from vignette studies. However, these studies are plagued by the fact that there is no guarantee that

respondents indeed report their true preferences, because neither their own money nor that of others is at stake. For example, there is no guarantee that respondents take the task seriously or that they do not give socially desirable responses. For these reasons, researchers began using experiments with real *monetary incentives*.

In many of these experiments, variations of the so-called *dictator game* (DG) are implemented (for overviews and interpretations, see Camerer, 2003; List, 2007; Bardsley, 2008). We briefly introduce this game here. In its classical form, the DG is a two-player game in which one of the players is assigned the role of the proposer (the 'dictator'), and the other player is the receiver. The proposer is given a certain money endowment E (e.g., 10 euros), and decides which fraction s of the endowment he or she wants to give to the receiver. The latter has only a passive role; he or she can only accept the gift. At the end of the game, the proposer earns $(1-s)E$, and the receiver earns sE. In the classic set-up, anonymity is preserved so that neither knows the identity of the other, and the game is played only once so that strategic considerations such as reciprocity do not play a role.

Konow (2000) adopted the standard DG and introduced the *third-party* DG. The experiment consists of two stages. In stage 1, all participants individually generate earnings in a real-effort task (preparing letters for mailing). Thereafter, participants are matched in pairs, and the sum of their earnings is credited to a joint account of the pair. In the second stage in one treatment ('standard dictator'), one subject of the pair is chosen to distribute the earned money between herself and her matched partner; in another treatment ('benevolent dictator'), a third party is chosen for this task. Importantly, the benevolent dictator's earning is independent of the allocation she implements. These two variations of the DG allow Konow (2000) to disentangle 'true' distributive justice principles (as expressed by the uninvolved benevolent dictator) from justice ideas that are intermingled with self-interest (as exhibited by the involved dictator). In a second treatment variation, Konow (2000) tests whether the support for the *Accountability Principle*, as observed in survey studies, carries over to situations in which real money is at stake. This is achieved by conducting two different versions of the first stage that differ in the way in which the real-effort task was rewarded. In the 'discretionary difference' treatment, each prepared

letter earned the same amount of money - and any differences in individual earnings came about through individual differences in productivity in letter preparation. In the 'exogenous difference' treatment, participants were given enough time such that everybody could produce the same number of letters. Differences in earning were generated by randomly assigning different per-letter rewards to the two players.

The reported results clearly support the accountability principle, and also show that allocation decisions when own stakes are involved are indeed strongly influenced by self-regarding concerns. More specifically, in the discretionary difference treatments, benevolent dictators almost always allocate the pair's joint earnings in proportion to the individuals' contribution in the real-effort task. In stark contrast, in the exogenous productivity difference treatment, benevolent dictators allocate the pair's joint earnings 50/50 - independent of the differences in individual earnings. In fact, almost 90% of the benevolent dictators allocated exactly equal shares. Standard dictators also take the accountability principle into account, and show a tendency to allocate joint earnings in proportion to individual earnings. However, the application of the principle is somewhat biased toward the self-interest of the dictator. Basically, all deviations from proportional allocations are in the direction favouring the dictator - and although allocations are significantly related to the discretionary input of the recipient, recipients receive only 30 cents more for every 100 cents more they contribute to the joint earnings. In the exogenous differences treatment, standard dictators allocate 50% or less to the recipient - and when they allocate in proportion to the arbitrary per-letter rewards, they do this when it favors them, indicating the effect of material self-interest.

Dickinson and Tiefenthaler (2002) used a similar third-party dictator experimental design to investigate the difference of fairness conceptions when dealing with *allocations* (inputs) or with *outcomes* (outputs). Whereas Konow (2000) implicitly induced a utility function that is linear in money and the same for everybody, these authors induce non-linear utility in money income that differs across participants. An important consequence of this variation is that equal allocations do not translate to equal money earnings. Similar to the Konow (2000) study, recipient-participants in one treatment earned their rights, while in another one this was not the case. On aggre-

gate, about 54% of benevolent dictators chose an allocation that equalizes the outcomes - whereas only about 4% chose an allocation with equal inputs (and unequal outputs). In addition, about 11% chose an allocation that maximized the joint outcome but led to unequal individual outcomes. When comparing the no-earned rights with the earned-rights treatment, Dickinson and Tiefenthaler observed a significant shift away from equal outcomes. While in the former case about 62% of all uninvolved dictators chose allocations that equalize outcomes, this percentage dropped to about 46% in the latter case. Hence, also in a non-linear (and therefore more complex) environment, equality of outcomes and the accountability principle seem to be important. An interesting side result of this study is that women seem to be less sensitive to the introduction of earned rights than men. Specifically, in the earned-rights treatment, 58% of women chose allocations equalizing outcomes (compared to only 35% of the male participants).

In a particularly rich study, Durante and Putterman (2009) investigate how the level of redistributive taxation set by a dictator depends on whether or not the dictator is personally affected by the tax. The results indicate that most dictators favour a more equal income distribution, especially when the pre-tax income distribution is based on luck rather than ability. However, when dictators are personally affected, the tax level they choose is strongly biased toward their self-interest. They choose a higher (lower) tax when they expect to be relatively poor (rich), and this effect is stronger when the income distribution is certain rather than when it is uncertain. Shayo and Harel (2012) report a related result. They find that subjects are more likely to vote for an allocation which is biased in their own favour and less likely to vote for an equitable allocation if it is more likely that their vote will be pivotal. Hence, choices seem to coincide less with those of a impartial third party when it is more likely that an alternative choice will be to their own advantage.

In a paper, building upon Konow (2000), Konow et al. (2009) empirically examine the possible determinants for equity and equality. They specifically investigate if and how the relative importance of equity and equality depends on *personal characteristics* and *interpersonal factors*. In line with earlier evidence, the authors find that in impersonal settings participants strongly favour outcomes consistent with equity (proportionality). This re-

sult is robust to variations in cultural (Japan and US) and demographic (age, income, work hours, race and gender) backgrounds of participants. Interestingly, however, introducing interpersonal factors and decreasing social distance has significant effects - in that it leads to shifting allocations from equity to equality. The authors conclude that social preferences are constructed by 'morals' and 'mores' - where the former refers to the moral preferences people have when they are in the role of a neutral non-involved arbitrator, and the latter refers to social preferences activated by personal considerations.

Distributive justice and earned rights

Redistribution usually does not take place in an idealistic societal vacuum. When pondering just distributions, people may take into account the fact that some positions embody some sort of 'right' or 'claim'. For instance, in discussions about pension reform, some may perceive that people belonging to the older generation have the right to receive a certain level of benefits. Such *rights and claims* are studied by Gächter and Riedl (2005) and Gächter and Riedl (2006). Pairs of participants acquire asymmetric monetary claims through a real-effort task. Thereafter, nature decides whether the claims are actually paid out or if the parties have to bargain over a smaller pie - where it is impossible to satisfy both claims simultaneously. The two participants and the impartial third parties are asked for their judgments regarding the just division of the reduced pie. Importantly, the claims are economically sunk. Nevertheless, the vast majority of both participants and third parties take these claims into account when formulating their judgment regarding fair distribution. Specifically, the distribution proportional to the acquired claims figures prominently in the proposed allocations. In addition, some preference for progressivity is observed, in that the proposed distributions become relatively more equal with increasing asymmetries in the claims.

Chavanne et al. (2009) utilize third-party dictator experiments to explore redistribution preferences in the presence of entitlements and inequalities. Specifically, in their set-up, one of two stakeholders is endowed with money and a third-party dictator can redistribute any portion of this endowment to the stakeholder without endowment. Hence, the third party has to ac-

tively take money away from one person to increase the earnings of another. The authors investigate how different ways of legitimizing the initial endowment alter the benevolent dictator's redistribution decision. In one pair of treatments, the endowed *position* is either assigned randomly or through the performance in a test. In another pair of treatments, the *amount* of endowment was either determined randomly or acquired by working on a word search task. The authors find that redistribution takes place - but that it depends on the way in which the endowed position and the endowment itself are received. Most redistribution takes place when the position and the endowment are randomly assigned. In this case, third parties (on average) equalize the earnings of stakeholders. When the endowment position or the amount of endowment is earned, only between 35% and 41% are redistributed to the party without endowment.

Distributive justice in the face of risk and uncertainty

Despite the prevalence of risk and uncertainty that accompany everyday life and economic activities, most of the surveyed studies on justice principles utilize a deterministic amount of income. A recent study, Cappelen et al. (2013), investigates fairness views about *risk taking*, and examine whether people's ideas regarding justice focus mainly on *ex-ante* opportunities or *ex-post* outcomes. The *ex-ante* view (focusing on initial opportunities) provides a fairness-based argument for no redistribution of eventual *ex-post* gains and losses. In contrast, the *ex-post* view (focusing on outcomes) provides a fairness-based rationale for eliminating *ex-post* inequalities coming from risky decisions. To experimentally investigate fairness views of risk taking, the authors implemented a two-stage design. In the first stage, participants had to choose between risky and safe alternatives. Participants in the second stage were paired, and earnings resulting from the first-stage decisions were pooled. Participants were then informed about choices and outcomes of the risk-taking stage, and had to distribute the pooled earnings. In addition, some participants acted as uninvolved third parties (spectators) who did not participate in the risk-taking task, and were asked to distribute the pooled income between the two involved parties (stakeholders). The authors report the following main results: (i) the majority of spectators distribute

total earnings equally; (ii) however, many participants did not deem it fair to equalize income when there is a difference in risk taking, but found it fair if the difference is in luck; (iii) the distribution decisions are independent of the costs of avoiding risk; and (iv) choices of spectators and stakeholders seem to reflect the same set of fairness considerations.

Summary

Two methods have been used to measure the moral preferences of individuals regarding income differences, while controlling for potential biases created by self-interest. One method is to put people behind a veil of ignorance. Studies show that people have a preference for maximizing the average income in society, subject to a floor constraint. This can interpreted as saying that people are quite willing to trade off some equality if this is compensated by extra efficiency. The second method consists of having people choose among income distributions over others as a third party dictator. The results of these studies suggest that full income equality is the normative ideal, and that the willingness to deviate from this ideal varies systematically with a number of elements in the decision environment. A major element is the reason that lies behind income inequalities. Many people in their role as third party seem to follow the accountability principle. Income inequalities that arise from factors beyond a person's control (luck and disability) should be repaired, while inequalities that are within a person's control (effort) are tolerated. Implementation of the principle, however, also depends on the context. The more 'social' the setting and the smaller the social distance, the higher the relative weight put on equality versus equity (proportionality). The relative weight on equality is also higher for women than for men.

2.3.2 Preferences regarding income distribution: in front of the veil ignorance

The previous section focused mainly on distribution games in which the allocator is not involved - in the sense that own earnings are not at stake when making the distribution decision. However, in most circumstances people know their positions, and it is thus likely that some tension exists between self-serving and social preferences. An influential early study in psychology

investigating such a situation is Loewenstein et al. (1989). They used the vignette method for eliciting the weight people put on their own income relative to the income of others. The authors implemented different scenarios where subjects had to imagine themselves of being one of two disputants. Subjects were then confronted with different dispute outcomes allocating more or less money to one of the two disputants and had to indicate their satisfaction with the proposed outcome on an 11-point Likert scale. The authors found that subjects showed dissatisfaction when they were behind the other disputant as well as when they were in an advantageous position. This study and its results can be seen as an important source of inspiration for experimental studies in economics as well as the development of social and other-regarding preferences models in economics (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000).

In situations where decisions affect the decision makers as well as other persons well-being, economic experiments offer a tool with which one can investigate the effect of different institutional environments in an incentivized and a controlled way. Many experimental setups build on the standard DG, as described in the previous section.

The standard DG was first implemented by Forsythe et al. (1994). They find that dictators on average decide to give about 1 of their 5 endowment to the receiver. Dozens of replications indicate that this is representative for the outcome of DGs (see Camerer, 2003). Typically, more than 60% of the subjects in the role of the allocator choose a positive transfer and the mean transfer amounts to approximately 20% of the endowment. At the same time, dictators' behaviour is very heterogeneous: there is a substantial fraction of dictators (about 35%) who give nothing to the receiver; another large fraction (25%) gives the receiver an equal share, while the rest of the dictators give amounts somewhere between these extremes.

The influence of the size of the stakes

One variation of the DG was introduced to examine if it matters whether the monetary stakes are *real* or *hypothetical*. Forsythe et al. (1994) compare dictator decisions for pie sizes of $5 in two treatments. In one, the DG was played with real monetary stakes; in the other, the stakes were merely hy-

pothetical. They find that the hypothetical decisions were more generous than the real ones, and reject the hypothesis that the distributions of proposals are the same in both treatments. Sefton (1992) and Krawczyk and Le Lec (2008) find similar results. The latter conclude, 'sharing equally in dictator game-like situations may be a socially-desirable norm of behaviour, which however is quite easily overridden when (sufficient) monetary incentives come into play'.

Is giving behaviour sensitive to the *size of the pie* to be distributed? Comparing two treatments capturing non-hypothetical decisions, with stakes of US$5 and US$10, Forsythe et al. (1994) find no significant effect on giving behaviour. However, the difference in pie sizes is only US$5. Carpenter et al. (2005) implement a larger difference of $90. They find that increasing the stakes from $10 to $100 has no statistically significant effect on behaviour in the DG. Similarly, List and Cherry (2008) find no significant difference between allocations comparing a DG with stakes of 20 and 100. Hence, it seems that the results from DGs are not an artefact of the relatively small stakes involved.

The process that generates income and decision power

Do distributional preferences depend on whether the initial endowments are earned or not, and do they depend on whether the role of the dictator is earned or randomly assigned?

Hoffman et al. (1994) report an experiment in which subjects could earn being in the advantage role of the dictator. Subjects first took part in a general knowledge quiz where those with the best performance were assigned the role of the dictator. The receivers in this contest-entitlement treatment ended up with a much lower payoff than those in the control treatment where the roles were allocated randomly as in the standard DG. Subjects among the top performers in the knowledge quiz seemingly felt they had earned their position and thus a property right over their initial endowment.

Jakiela (2009) reports the results of a comparison between a standard DG and the 'taking game', where the dictator's partner holds the whole endowment in the beginning. She finds that dictators allocate themselves a larger share when they themselves are endowed, with the endowment being

determined by luck, than when their partner is endowed with the money and reallocation means to actively take money away. In an additional set of treatments, the standard DG was preceded by a piece-rate effort task (sorting dried beans out of a bucket) that determined the subjects' endowments. The author finds that subjects allocate more to themselves when they earn their endowment compared to when they win it. Similarly, Oxoby and Spraggon (2008) report that in the standard DG dictators allocate on average 20% to the receivers, whereas when the dictator had earned the wealth, transfers were close to zero. On the other hand, if the receiver had earned the wealth, dictators sometimes even gave more than 50% of the pie to the receiver. This suggests that legitimizing of assets creates property rights that participants tend to respect, regardless of whether the powerful or the powerless accumulate these rights (see also Ruffle, 1998; List, 2007; Krawczyk and Le Lec, 2008; List and Cherry, 2008; Durante and Putterman, 2009).

Is the earnings-based notion of justice as distinctive, if high productivity is mainly due to pure talent than if it is due to the effort that is put into production? The experiment reported in Cappelen et al. (2007) is informative with respect to that question. As in the experiments reviewed above, the distribution phase is preceded by a production phase where subjects are asked to choose how much of their endowment to invest in two different games. An exogenously given rate of return determines each player's eventual contribution - those with a high rate of return would quadruple their effort investment, whereas those with a low rate of return would merely double it. In the distribution phase, subjects are paired with players differing with respect to their rate of return for the two games, are informed about the opponent's investment, rate of return and the total contribution. They are then asked to decide about how to distribute the total income like in a conventional DG game. The results show that many participants distinguish between factors that are within subject's control (investment/effort) and those that are exogenous (rate of return/talent) in the sense that they only perceive inequalities due to factors within individual control as justifiable.

Social identity and social distance

Other factors that are found to affect distributional preferences are *social identity* and *social distance*. Hoffman et al. (1994) employed a so-called double-blind procedure that guaranteed complete anonymity, in the sense that neither the experimenter nor the other subjects could observe a subject's decision and payoff. The authors find that under such a strict anonymity setting, a majority of dictators (64%) give nothing to the receiver, while in Forsythe et al. (1994) only 36% give nothing (see also Hoffman et al., 1996).

Charness and Gneezy (2008) examine the opposite effect of decreasing social distance on giving behaviour in a DG by comparing behaviour in the classic DG approach with a treatment in which participants knew the family name of the subject they were matched with. When the names were known, dictators were significantly more generous and allocated a higher portion to the receiver (see also Johannesson and Persson, 2000). Recently, Leider et al. (2009), D'Excelle and Riedl (2010) and Goeree et al. (2010) investigated dictator giving behaviour in real existing social networks of Harvard undergraduates, female high school students and household heads of a village in rural Nicaragua, respectively. The authors mapped the friendship network as well as other social and economic links (D'Excelle and Riedl, 2010), which makes it possible to calculate the social distance between any two people in the network. Thereafter, people in the investigated networks participated in a series of DGs, with some being dictators and others receivers. In all three studies, dictator giving significantly decreased with larger social distances.

Klor and Shayo (2010) study the effects of social identity (group membership) on voting over redistribution. Subjects were divided into two groups according to their field of study. They were randomly assigned different income levels and were informed about their own income, the overall mean income and the mean income of each group. Thereafter, they voted anonymously over a redistributive tax regime that was determined by majority rule. The tax revenue would then be equally distributed among all subjects. This procedure was repeated 40 times without giving subjects information about the effective tax rate and their individual payoff after each round. Comparing subject's behaviour to a treatment in which subjects did not know about the group assignment, the authors find that identification with a group indeed

affects redistribution preferences. More than a third of the subjects (most of them facing a cost of opting for the well-being of their group that was not too high) did not maximize their payoff, but chose the tax rate that was best for the average member of their group. The authors can exclude other motives, e.g. efficiency concerns or inequality aversion as being accountable for the observed behavioural differences.

Gender, age and ideology

As mentioned before, there is a substantial degree of *heterogeneity* in social preferences and some observable individual characteristics seem to differentiate those who give more from those who give less. Several papers allude to the relationship of gender and giving behaviour (for a recent survey, see Croson and Gneezy, 2009). Eckel and Grossman (1998) report the results of a double-blind DG. They find that women are more generous than men: on average, men give half of what women give to their anonymous partner. Bolton and Katok (1995), in contrast, find no gender difference when investigating dictator's choices applying only subject-subject anonymity. Cox and Deck (2006) compare behaviour across genders in allocation decisions and conclude that behavioural differences between men and women are context dependent. Women tend to be more generous than men when social distance is low (social separation between the subject and all other people who are present for the experiment), monetary cost of generosity is low (forgone amount of money when subject chooses a generous action), and when there is an absence of reciprocal motivation (as in the DG).

Apart from the effect of gender, Bellemare et al. (2008) find that older people have a stronger preference for income equality than younger people (< 35 years).

Perhaps unsurprisingly, ideological orientations are also strongly related to redistribution preferences. Esarey et al. (2009) find that survey measures of individuals' economic ideologies can predict their preferences for redistribution programs that combine income equalization and social insurance. In the first stage of their experiment, an individual production task determined each subject's endowment. In the second stage, each individual within one treatment faced the same probability of losing 80% of the endowment. Sub-

jects were asked to vote on an income redistribution plan, a tax rate between 0 and 100%, which would be deduced from their incomes *before* the potential occurrence of the random shock, where the median of the choices became the effective tax rate for the following periods. The tax revenue would then be equally distributed among all subjects. More economically liberal subjects (as assessed with a questionnaire) voted for higher tax rates than the more economically conservative ones - however, only in the treatment with a moderate risk of a random shock. The authors interpret this as liberals acting in accordance with the idea that individuals should be protected from bad luck, while conservatives act in accordance with the idea that bad luck is 'something to be suffered and good luck [...] something to be enjoyed' (Esarey et al., 2009, p. 5).

The role of institutions : markets and politics

In most of the experiments reviewed above, the decision maker (dictator) has absolute power over the income distribution. The advantage of such a setup is that it gives a very clear and direct view on people's social preferences. In reality, of course, distributional outcomes take shape in a much richer institutional context, which may constrain or facilitate the intensity of revealed social preferences.

Experiments have shown, for example, that market competition can be an important check on the role of distributional preferences (Fehr and Schmidt, 1999). Even if all players on one side of a market prefer an equitable outcome, the competition between them might still lead to quite an unequal result. The reason is that an individual player has no control over the outcome, and that coordination is usually difficult to achieve. Interestingly, the reverse may also be true. In some important circumstances the presence of social preferences can nullify the impact of competition on market outcomes (Fehr and Falk, 1999). For instance, wage cuts are rarely observed even in times of high unemployment, because managers fear that employees may respond with less effort and more on-the-job consumption.

Also political institutions may interact in intricate ways with distributional preferences, as some experimental studies have shown. Tyran and Sausgruber (2006) experimentally study the effect of voting on redistribu-

tion. They artificially create poor, middle class and rich subjects by giving them different initial endowments. While classical theory assuming narrow self-interest predicts that only the poor would vote for redistribution, the authors find that, next to the poor, also 70% of the middle class and even one-third of the rich voted for redistribution from the rich to the poor. Cabrales et al. (2006) find a seemingly opposite result, namely that majority voting does not lead to redistribution. An important difference with the previous study, however, is that differences in income are not just random and exogenous, but are partly endogenous and determined by the costly effort individuals exert. Clearly, this reduces the willingness of the rich to vote for redistribution. In a related study Höchtl et al. (2012) show that the structure of income classes is decisive for the relevance of fairness preferences for majority voting on redistribution. The relative size of the classes (poor versus rich) and which of the two is in majority determines whether fair-minded voters are pivotal, which in turn affects redistribution outcomes. It is found that a given level of fairness concerns matters for the aggregate outcome when the majority is rich, whereas redistribution outcomes seem to be motivated by selfish-interest when the majority is poor.

Two other studies experimentally investigate the interaction between political institutions and social preferences. Messer et al. (2010) study the provision of public goods via public referenda. They find that individuals' preferences for social efficiency lead to deviations from the selfish voting outcome in the direction of a higher likelihood of implementation of welfare increasing outcomes. Paetzel et al. (2012) investigate how social preferences affect voting for a reform that increases total income but at the same time also increases inequality among voters. They find that a considerable share of voters among 'reform losers' vote for the good of society rather than their own pocketbook. These voters outweigh the share of 'reform winners' with a preference for equality, so that in the aggregate voting outcomes tend to be in favour of the reform.

2.3.3 Strategic interaction and redistribution

Often, redistribution also involves a strategic element. When the decision about sharing risks is made *before* uncertainty about the individual out-

comes is resolved, individuals may beforehand agree that the lucky should support the unlucky. When the risk has materialized, however, the lucky may have an incentive to reconsider the agreement. Therefore, in the absence of enforceable contracts, voluntary risk sharing is akin to a social dilemma. It is in the players' joint interest that everyone cooperates and sticks to the agreement, but individual players may have an incentive to defect and renegotiate. Several experiments have examined how people resolve the conflict between joint interest and self-interest. Are people willing to cooperate, to share risks, or do they take a 'free ride' whenever they can? Which factors determine whether a cooperative outcome is attainable? What does this tell us about people's social preferences?

Risk sharing and insurance games

Selten and Ockenfels (1998) introduced the so-called *Solidarity Game*, which offers a basic set-up for investigating redistribution preferences when people are exposed to risk. Each of three players has a chance of 2/3 to receive an income of 10 DM, and a chance of 1/3 to receive 0 DM. Before the players know whether they receive 10 DM or 0 DM, each player is asked how much he or she is willing to give to a player who receives 0 DM in the event that he or she receives 10 DM. Thus, subjects can share risks *ex-post* here, but there is no strategic element involved since it is a one-shot game. Are the winners willing to compensate the losers, and how does this depend on the number of losers? The results indicate that 79% of the winners are willing to transfer a positive amount to the loser(s). Remarkably, for 50% of the winners the total amount they were willing to transfer (about 3 DM) did not depend on whether there were one or two losers. This implies that a single loser would receive a total transfer of 6 DM (3 DM from each winner), leading to a very equitable income distribution (7-7-6) - whereas two losers would each receive only 1.5 DM, leading to a very skewed income distribution (7-1.5-1.5).

Charness and Genicot (2009) experimentally investigate a *voluntary risk-sharing* game. In the experiment a subject is matched with another subject and each is endowed with a fixed income. Additionally, in each period it is randomly determined which of the two subjects in a pair would get an extra amount on top. After both outcomes are observed, subjects can choose to

transfer money to the other subject. This part of the game is similar to a DG. One difference is that not only the rich but also the poor players can make a transfer. Additionally, there is a strategic element included since the game is played repeatedly and subjects stay with their partner for an uncertain number of periods based on a certain continuation probability after each period. After that subjects are matched with a new partner and play the game again. The game segments vary with respect to whether the fixed incomes are equal or unequal. Finally, one of the periods is randomly determined to be relevant for payment. The authors find that subjects do share risks, with higher transfers coming from subjects who got the extra amount. However, also the other side often makes a small transfer - the authors speculate that this might be a signal of intent. They also find evidence for reciprocal behaviour in the sense that a subject's transfer is higher, the higher the first transfer made by his partner. The more risk-averse subjects are and the higher the continuation probability, the higher the level of risk sharing. Inequality in fixed incomes instead leads to a decrease in risk sharing.

Charness and Genicot (2009) investigate risk sharing without commitment, whereas in Barr and Genicot (2008) the level of commitment *is varied*. The authors conduct a field experiment in Zimbabwe: First subjects have the choice between six gambles varying in average return and riskiness. In round 1 of the experiment subjects play this gamble choice game individually, the possibility of risk sharing being excluded. Before taking part in round 2, subjects are invited to form risk sharing groups, implying that all members of one sharing group would pool the money they won in the gambles and distribute it equally among all group members. In two of the treatments, once a subject decides to opt into the collective insurance, there is still the possibility to opt-out, either in public or private, *after* a subject's personal outcome is observed. However, in a third treatment subjects face full commitment, so if they decide to join a sharing group this is an effective decision and they *do not* have the possibility to change their mind when being informed about their individual outcome. The authors find that subjects in the latter treatment are more likely to form risk sharing groups. Only 31% do not join a risk sharing group whereas in the other two treatments about 60% do not join. Additionally, subjects in the full commitment treatment take more risks in the gamble choice game. On average the groups

formed in that treatment are larger and include 6.9 people compared with 6.5 in the treatment with private defection and four in that with public defection.

Chaudhuri et al. (2010) investigate risk sharing in groups (5 vs. 25 members) that play together for at least 20 periods but face uncertainty about the exact number of periods. In each period subjects are first informed about the outcome of a random draw that determines their endowment for this period - they either get a high or low endowment. Additionally, they get to know how many other people in the group received a high endowment. Then they are asked how much money they want to place in a group account that would be equally distributed among all members of the group. This so-called *insurance game* is a game of collective action with heterogeneous endowments among subjects. When making their decision about how much to put into the group account in a certain period, subjects face the uncertainty about how much other group members will actually contribute. Repeated play may trigger strategic considerations based on the expectation of reciprocal behaviour since subjects are uncertain about their endowments in the following periods. The number of group members with a high or low endowment may differ between periods. The authors find that in small groups contributions to the pool are significantly higher compared with large groups, but that there is no complete risk sharing. In groups that are self-selected, by requiring subjects to register for the experiment as a group, risk sharing is significantly higher than in non self-selected groups.

Overlapping generations and intergenerational transfers

A temporal structure particularly relevant in the domain of pensions and health insurance is that of a sequence of overlapping generations. For example, in a pay-as- you-go (PAYG) system the currently retired generation is supported by the currently working generation; when the latter generation retires, they will be supported by the next generation (and so on). Such a system of intergenerational transfers, however, may suffer from a temporal credibility problem. What is the guarantee that the currently working generation will receive the same level of support from the next generation, once they retire? Every working generation may experience an incentive to recon-

sider the level of support to the currently old generation. One behavioural mechanism that could make a PAYG system self-enforcing is that of intergenerational reciprocity. The present generation receives support in relation to the support they gave to the previous generation (Hammond, 1975; Kotlikoff et al., 1988).

Van der Heijden et al. (1998) use experiments to examine the relevance of such cross-generational reciprocity (see also Offerman et al., 2001). They employ a simple overlapping-generations game that abstracts from all complexities that could blur the view on this central idea. The game consists of a sequence of players (generations). Each player lives for two periods. In the first period, the (young) player has a high income; in the second period, the (old) player has a low income. Players cannot save, so that efficient income smoothing is possible only through intergenerational transfers. Player (generation) P_t decides on the transfer (pension) T_t to player P_{t-1} ; player P_{t+1} decides on the transfer T_{t+1} to player P_t ; player P_{t+2} decides on the transfer T_{t+2} to player P_{t+1}, and so on. The experiment examines whether there is a positive relationship between T_{t+1} and T_t. Is the transfer that a player receives from the next player related to the transfer that this player gave to the previous player? Moreover, the paper examines whether such intergenerational reciprocity increases the viability of a PAYG transfer scheme. For that purpose, two information treatments are implemented. In one treatment, a player knows T_{t-1} when deciding upon T_t. In the other treatment, a player does not know T_{t-1} when deciding upon T_t. Obviously, the latter treatment rules out any role for monitoring and reciprocity.

The results of the experiment are clear. There is no evidence whatsoever for intergenerational reciprocity. The level of the transfer in period t+1 is unrelated to the level of the transfer in period t. Player P_t is neither rewarded nor punished by player P_{t+1} for the way he or she treated player P_{t-1}. This result is corroborated by the finding that the average level of transfers is the same in the two information treatments. It makes no difference whether or not the previous levels of transfers can be observed. Still, the average level of transfers can be considered quite high. The payoffs in the game would be equal to nine without transfers (individual rationality), and 25 with optimal transfers (collective rationality). With the observed level of transfers, the realized average payoff is 21. So, it might be said that a fairly efficient vol-

untary pension system emerges. This is quite remarkable, in view of the fact that no commitment possibilities are available. A standard game theoretical analysis based on purely selfish agents would predict no transfers at all.

Some extensions of this pension game have been studied experimentally. One of these allows for private (retirement) savings besides the option to use intergenerational transfers for that purpose (Van der Heijden et al., 1998). The results show that the possibility of individual savings erodes the support for intergenerational transfers. This occurs, despite the fact that - in the experiment - intergenerational transfers are more efficient than private savings. The main attraction of private savings in comparison with a PAYG system is that the former suffer none of the uncertainty of the latter that the system will be maintained to the same degree in the future.

Güth et al. (2002) study an overlapping-generations experiment with multiple 'families' in which two types of intergenerational transfers are possible. A generation can make voluntary transfers (S_t) to the previous generation (essentially a PAYG pension to their parents). In addition, a generation can make a transfer (G_t) to the next generation (essentially an investment in the human capital and, thus, the earnings potential of their children). One of the aims of the experiment is to investigate the relationship between S_{t+1} and S_t, as well as the relationship between S_{t+1} and G_t. In other words, is a generation (when old) rewarded for how it treated its parents and/or for how it treated its children? The experimental results suggest that in fact both types of relationships are rather weak. Again, reciprocity - direct or indirect - does not seem to be a major factor in explaining the support for intergenerational transfers.

2.4 The main lessons

Principles of redistributive justice are utilitarian with a floor constraint

People share certain principles of redistributive justice when they are *behind a veil of ignorance*. They are willing to trade off some inequality for some efficiency. They neither want to implement the Rawlsian rule that makes the potentially worst-off better-off, disregarding efficiency, nor do

they want to stick to pure utilitarianism (Harsanyi, 1955), which maximizes total utility independent of the distribution of individual well-being. Rather, the most preferred rule of justice is utilitarianism combined with a safety net for the poorest. In other words, people find it acceptable that some individuals are worse-off - as long as they are not too disadvantaged - and if this is compensated by a larger number of other people being better-off.

People are averse to inequality, but this aversion varies with the source of inequality

In symmetric situations in which people do not differ from each other in important aspects, the equal division or equal sharing norm is prevalent. In asymmetric situations, shared distribution norms seem also to exist - even if these norms lead to inequality. People thus seem quite tolerant of inequality under certain conditions. The acceptance of inequality strongly depends on the source of inequality. Accountability (Konow, 1996) and equity - in the sense of proportionality - are the leading principles. Income inequalities are acceptable when they can be traced back to factors within people's control - but not if they are the result of factors beyond their control.

Social preferences are relevant even if the veil of ignorance has been lifted

Experiments have shown not only that people share justice ideas when their own material well-being is not at stake, but that they care for the well-being of others and for the aggregate outcome even if it comes at material cost to them. People leave money on the table for anonymous others, even if they could easily get away with taking everything.

Social preferences display a self-serving bias

If people know their own position in society, preferences for redistribution are strongly coloured by self-interest. This is especially important in situations in which people are not symmetric. While people easily agree that equality is a good fairness norm when everybody is equal in all of the aspects deemed important, people tend to disagree on the fairness norm when they

differ with regard to important characteristics. For instance, in asymmetric situations, disadvantaged people tend to favour equality, whereas the advantaged propagate proportionality. Moreover, individuals in an advantaged position typically have a different perspective on accountability than do those in a disadvantaged position. What is deemed to be within or beyond a person's range of control varies across individuals, depending on their own interests on the matter. This is reminiscent of what psychologists call 'attribution bias', according to which people tend to claim successes as being due to merit, while explaining failure as a result of bad fortune.

Social preferences depend on the income-generation process

Accountability is not only important for redistributive justice but also shapes preferences for redistribution in front of the veil of ignorance. People are much more willing to redistribute income at a personal cost when they feel that the recipient deserves it. It is easier to accept redistribution in favour of low-income earners whose low income is due to bad luck than when the low income is due to low individual efforts. Similarly, redistribution towards less productive people is more easily accepted if the low productivity is beyond one's responsibility.

Social preferences are heterogeneous

Although social preferences are ubiquitous, not all people reveal social preferences. In addition, those who reveal social preferences do show significant variation in how strongly they take the well-being of others into account. Much of this observed heterogeneity is still unexplained, but a few personal characteristics show significant correlation with expressed social preferences. Women seem to be more generous than men, but their generosity is also more sensitive with respect to environmental specifics. Furthermore, older people seem to be more sensitive regarding income inequalities than are younger people. Real and perceived social distance between the persons involved in the redistribution also explains parts of the variation in expressed social preferences, with preferences for redistribution increasing with decreasing social distance.

Effect of political and economic institutions is ambiguous

Market competition constrains the impact of social preferences on outcomes - but the reverse is also true. Whether political institutions constrain or facilitate the impact of social preferences is largely unexplored. What existing studies have shown is that political institutions interact with social preferences in a non-trivial way, and that the specifics of the setting may tip the impact one way or another.

Social preferences are fragile

Social preferences are not only heterogeneous; their expression is also sensitive to institutional specifics and to beliefs about the social preferences of others. The willingness to redistribute income is sometimes influenced by economically unimportant details of the decision environment. In addition, generosity and cooperation is often conditional - in the sense that it is only expressed if people believe that others are also generous and cooperative. This implies an important role for expectations and trust for the support of redistribution schemes.

Social preferences across generations do not rely on reciprocity

The experimental evidence on altruism and social preferences across generations delivers a clear and to some extent surprising picture. There is no evidence for intergenerational reciprocity, in the sense that a generation that received support from the previous generation is more likely to support the next. Subsequent generations supported each other more or less unconditionally. Of course, this lesson is based on the supposition that this (strategic) component of the interaction between experimental generations is representative for those of 'real' generations. For example, in the experiments, the different 'generations' are all from the same age cohort (they are all students); whereas 'real' generations are obviously from different cohorts.

2.5 Perspectives for further experimental research

The experimental evidence we have surveyed convincingly shows that most people take into account their own justice principles or the perceived fairness ideas of others when deciding on distributive tasks - be these decisions taken in solitary circumstances or taken in situations where they have to interact with others. Pensions and social insurances are inherently (re) distributive, and this evidence is therefore important when one wishes to discuss individual and political pension and insurance options in an informed way. What is largely missing in the experimental designs is some reflection of the fact that (re)distributional decisions have *time dimension* and are prone to *risk* and *uncertainty*. These, time, risk and uncertainty, perspectives are particularly important for pensions and social insurance.

Therefore, a first set of research questions may tackle issues regarding principles of justice when the consequences take immediate effect (or only with some delay) - and only affect the present generation (or also later generations). The surveyed studies have shown that justice principles allow for inequalities if they can be linked to circumstances for which a person can be made accountable. If we translate this to the pension problem, then the idea of accountability implies that people may be willing to accept that others receive higher pension payments if, for instance, this is based on higher productivity due to training followed - but will be less willing to do so if the higher productivity is based on pure talent or luck. An important complicating factor with pensions is that these principles affect not (only) one's own generation but (also) other generations. The sustainability of a pension system based on intergenerational solidarity calls for both an extension of justice principles across generations and solidarity between different social classes within a generation. Not much is known about the fairness ideas of people in such situations. How should the benefits be distributed between different income classes and across generations? Perhaps even more important is the question of how the *burden* in times of distress should be distributed between generations and social classes. A similar quandary applies to social insurance, where the benefits and costs have to be distributed between people with different income (potential) and different risks. A first small step in analyzing fairness ideas in such contexts was taken by Cap-

pelen et al. (2013) in their study of the fairness perceptions of risk-taking. However, many questions remain: What is the fairness perception of the trade-off between risk-taking and income? To what extent should a person who deliberately took a high risk and earned good income through good luck be made accountable for the good income? Should she be treated differently from a person who opted for low risk and had bad luck? These are important questions at the heart of any social insurance scheme.

The above briefly discusses possible research into the normative basis of pension and social insurance schemes when people place themselves outside the scheme. In reality, however, people are often stakeholders in such schemes. The research on distribution problems clearly indicates that the distribution decisions of stakeholders are influenced by the trade-off between fairness and material self-interest. A similar trade-off is to be expected when it comes to decisions about the distribution of benefits and costs regarding pension and social insurance schemes. Having established the underlying normative foundations in distributional decisions in which the time and risk dimension play an important role, the logical next step involves investigating the effect of personal involvement, in the context of seeking to know more about the sustainability of particular pension and social security schemes. Interesting and important variations and extensions of experimental designs come again from the very nature of pensions and social insurance. For instance, self-interest may also influence behaviour via historically grown entitlements, a phenomenon observed in experimental bargaining and negotiations, but not yet experimentally examined in the context of redistribution between different income classes and generations.

In democracies, decisions about the redistributive consequences of pensions and social insurance are not implemented dictatorially but via a political process. Therefore, it is important to extend the small body of literature on the political economy of (re)distribution to the area of pensions and insurance. In future, it will be important to extend this aspect into an inter-generational setting with and without the involvement of risk and uncertainty. Building on experimental research into the justice principles regarding (re)distribution within and between generations, one can design institutions that maximize political support for sustaining and/or reforming economically meaningful pension and social insurance systems.

2.6 Conclusions and implications

This paper surveys the experimental evidence that deals with a major constituent element of solidarity: redistribution. While the evidence clearly shows that people share some basic willingness to support redistribution in general, the evidence also points to the limits to this support, which are influenced by various factors such as the source of inequality, social and personal characteristics and the institutional environment in which such redistribution takes place. The structure and distributional consequences of solidarity-based pension schemes have to be in line with generally shared fairness norms, and must take into account their limits. Otherwise, these schemes will lose societal support, and open the door to a host of adverse consequences.

An important message is that inequalities among people with unequal characteristics are acceptable to a large majority as long as there are good reasons for these inequalities, and if this acceptance does not lead to inefficiencies. In particular, the support for redistribution depends crucially on the sources of the inequality. One could argue that these distributional preferences reflect the possible disincentive effects of full insurance against all income risk, such as the incidence of income insurance on effort supply and other moral hazard effects. This requires that individuals bear at least part of the adverse consequences of their choices. It implies that if one wants to organize support for solidarity and redistribution, it is probably much less effective to emphasize the fact that the recipients *need* the support than it is to stress that they *deserve* the support. Importantly, such emphasis has to be transparent, because social preferences - and therefore support for re-distributional schemes implied in pension systems - have been shown to be fragile with respect to perceived injustices. Such transparent policies are also important because people's tendency to apply justice principles in a self-serving way may undermine solidarity when there is too much room left for idiosyncratic interpretations.

It has been suggested that the sustainability of collective pension schemes can be furthered by increasing the actuarial fairness and by reducing the (ex-ante) redistribution embodied in the system. One proposal is to make contributions dependent on observable risk characteristics. Experimental

evidence suggests, however, that such a proposal will probably not meet with much popular support if the characteristics involved are beyond a person's control (such as gender or age). There is likely to be much more approval if the differentiation is based on characteristics that can be reasonably expected to be due to a person's free choice (such as career decisions and having a partner or not). Generally, one may say that a collective pension scheme that reflects proportionality of benefits relative to the provided inputs, and that takes into account a person's accountability for his or her choices, will enjoy relatively strong support from the population. There is, however, a caveat to be made.

Although political and ideological differences usually do not lead to controversies about the underlying *principles* of fairness and justice, such controversies surface when it comes to the *interpretation* of these principles. Both left and right, rich and poor, men and women, Europeans and Americans by and large agree that people should suffer the consequences (and enjoy the fruits) of outcomes for which they can be held accountable. However, as soon as one starts trying to define for what precisely people can be held accountable, disagreement starts. For example, some will argue that talents and capacities are due to merit and education, or perhaps are a gift from God; others may perceive them as merely due to chance. Similarly, the rich may argue that their wealth is due to their own efforts and to the risks they have taken during their life, while the poor may claim that their low wealth level is largely due to bad luck. However, a number of virtually indisputable characteristics might form a basis upon which proportionality and accountability within a pension system could be based. For instance, few people will argue that individuals should be held accountable for their age or their longevity. This may partly explain the unfailing support for pension schemes, in general, and for intergenerational solidarity, in particular.

An important issue uncovered in some of the surveyed research is that entitlements (or moral property rights) can strongly shape the perception of fair distributions. Existing pension systems also create such entitlements, which make it difficult to implement necessary reforms in the face of current and future financial distress. The political conflicts surrounding efforts to increase the retirement age reflect how strongly entitlements can be perceived. Whether or not such entitlements are justified, policies targeting

reforms that change such entitlements have to take them seriously into account.

Experiments also indicate that the strength of social preferences is decreasing with social distance. If people do not feel that others belong to the same group in one way or another, they are less likely to feel responsible for their well-being. The support for redistribution via pension systems is crucially dependent on a sense of shared identity. This suggests that there is an upper bound on the level and scale at which solidarity and risk sharing can be organized. Even though efficiency and economies of scale may sometimes dictate that risk pooling be organized at a high level of aggregation, the support for the redistribution that such risk sharing entails is likely to decrease at higher levels of aggregation, especially in times of hardship.

Finally, we have seen that mutually beneficial voluntary risk sharing does occur - but often fails to reach efficient levels, even when the conditions seem relatively favourable. While this failure seems partly due to bounded rationality, another important element is lack of trust. This trust is fostered in a number of ways. One of these factors is the shadow of the future; another is the absence of outside options. Recent developments on the labour market, however, may erode both of these factors. In particular, increased mobility on the labour market may well erode employees' sense of identification with their employer and with their colleagues - and also decrease the period of time that employees are in the same pension fund. From this perspective, it is quite understandable that the support for solidarity is under stress. The upshot, in short, is that solidarity must be organized - even when there is broad consensus on the underlying principles of fairness and distributive justice.

Chapter 3

An Experimental Investigation of Risk Sharing and Adverse Selection*

*This chapter is joint work with Jan Potters and Arno Riedl and will be published in a special issue of the *Journal of Risk and Uncertainty* in April 2014. We are grateful to the associate editor of the journal and an anonymous referee for very helpful comments.

3.1 Introduction

Adverse selection can undermine the viability of voluntary social insurance and risk sharing arrangements. In case individuals with the most favorable risk profile are unwilling to join a risk pool, those with the next most favorable risk profile will also be unwilling to join and so on. In the end, only those individuals with the least attractive risk profile will be part of the risk sharing arrangement. This mechanism of (unraveling) adverse selection is often advanced as an argument in favor of mandatory participation in social insurance and similar risk sharing schemes.

Despite its key role in the economics of insurance and risk sharing, the empirical evidence for the force of adverse selection is not particularly strong. Fenger (2009) studies some European countries in which individuals have attained the opportunity to limit their participation in welfare schemes, such as health policies, unemployment policies and pension schemes. He concludes that there is no convincing evidence for adverse selection. The results for private insurance markets are more ambiguous, with evidence for adverse selection in some markets (health insurance, annuities), but not in others (life insurance, long term care). It is not clear why the evidence is mixed. One possibility is that very risk averse individuals not only have a higher demand for insurance, but may also behave more cautiously. A favorable risk profile may then go hand-in-hand with a high insurance coverage (Cutler et al., 2008). It may also be that individuals do not really know how their risk profile compares to that of others, in which case the scope for adverse selection is seriously reduced (Cohen and Siegelman, 2010).

In this paper we use laboratory experiments to investigate the behavioral relevance of adverse selection. A major advantage of the experimental approach is randomization. Decision makers can be randomly assigned to risk profiles thus ruling out any endogeneity bias. Another advantage is that participants can be informed about how their risk profile compares to that of others, so that this cannot be an obstacle for adverse selection. Even under these circumstances, however, it is still an open question whether selection effects will actually occur. Adverse selection places rather strong demands on the rationality of individuals. It relies on their ex-ante ability to predict whether a risk sharing arrangement will lead to favorable redis-

tribution ex-post. To do so requires them to anticipate the risk sharing decisions of others, and to adjust their own decisions accordingly. Moreover, the literature on social preferences (e.g. Fehr and Schmidt, 1999) suggests that even participants with a favorable risk profile may be willing to subsidize, in expectation, those with less favorable prospects.

In our experiment we implement a setup in which three individuals simultaneously decide whether they want to share their risks. Risks are simple two-outcome lotteries, and realizations are independent across players. Individuals make sharing decisions before the risks materialize. Those individuals who decide to share their risks divide (proportionally) the sum of the realized incomes among themselves. Across treatments we implement different forms of risk heterogeneity by varying the two outcomes of the lotteries and/or the corresponding probabilities. Lotteries can vary only in mean outcomes (leading to first-order stochastic dominance relations), vary only in spread (leading to mean-preserving spreads relations), or vary in both mean and spread (leading to a simple compensated risk increase). Risk profiles are common knowledge.

Our results display strong evidence for adverse selection in settings in which risk profiles can be ranked on the basis of first-order stochastic dominance. Individuals with a good risk are much less willing to enter a risk sharing arrangement than individuals with a bad risk. In addition, there is clear evidence for strategic adverse selection, which leads to unraveling. Individuals with an intermediate risk anticipate adverse selection of good risks and are in consequence less willing to share risk in this heterogeneous setting compared to a setting in which all individuals have the same (intermediate) risks. At the same time, our results display little evidence for adverse selection in case risks can only be ranked according to lottery spreads. When the bad risk is a mean-preserving spread of the good risk, this is no cause for adverse selection. Hence, what mainly seems to matter for adverse selection is differences in mean, not differences in spread.

Two other results are noteworthy. First, even though the observed comparative statics coincide with the (calibrated) predictions, the level of adverse selection is much less pronounced than might be expected. Second, even in case risk profiles are identical, so that adverse selection cannot be an issue, the overall level of voluntary risk sharing is lower than predicted.

We are not the first who experimentally study risk sharing, but to the best of our knowledge we are the first to study adverse selection in risk sharing. One of the earliest studies on risk sharing is Selten and Ockenfels (1998), who implement a three-player setting in which all players have the same probability of winning some income. Before the risks materialize, players decide simultaneously how much they will give to the losing player(s) in case they win. Their results indicate that a majority of the winners are willing to transfer money to the loser(s). Several papers build on this so-called solidarity game and study, for instance, the role of group size (Chaudhuri et al., 2010) or the role of reciprocity in a repeated setting (Charness and Genicot, 2009). An important feature of these papers is that the transfer decision is made conditional on being a winner. In this sense, risk sharing is ex-post, which is an important difference with our design where risks are shared ex-ante and decisions cannot be conditioned on the realization of risks.

Barr and Genicot (2008) conduct a field experiment in Zimbabwe where individuals can choose whether or not to share their stochastic income with other members of a group before risks are realized. An important difference with our paper is that risk profiles are endogenous: subjects choose from a set of lotteries before they decide about risk sharing. Moreover, the authors focus on the effect of commitment and examine what happens to the level of risk sharing if participants have the option to opt out after the incomes are realized. This is quite different from our paper, which focuses on adverse selection - that is, not participating in risk sharing before incomes are realized.

3.2 Design and Procedures

We implement four treatments that differ with respect to the risks that are assigned within sharing groups. Individuals are informed about their own and others' risk exposure, but face uncertainty about the actual outcome of the risks when making their decision about how much risk to share with their group members. This depicts an intermediate position between what John Rawls calls an individual's 'original position' behind the veil of ignorance and that beyond the veil (see Rawls, 1971). We first describe the

lottery assignment process and point out treatment differences. We then proceed with describing the risk sharing procedure and the experimental procedures including the elicitation of risk preferences.

3.2.1 Lottery Assignment

At the beginning of a session each participant is randomly assigned to one of the four treatments and a matching group with two other participants. Each of them faces a lottery that eventually results in a high income Y_i^H with probability p_i or a low income Y_i^l with probability $(1-p_i)$. The exact specification of the parameters depends on the treatment and which of three possible lottery types an individual is assigned to.

Table 1 gives an overview of the different lotteries separated by treatment. It specifies p_i, Y_i^H, Y_i^l, as well as the expected lottery value μ and the lottery spread δ. The design of our treatments allows the investigation of voluntary risk sharing under different forms of risk heterogeneity. In particular, we can observe to what extent the occurrence of selection effects hampers the effectiveness of risk sharing and to what extent risk sharing is reduced compared to a situation where risks are homogeneous.

Table 1 Treatments

HO	p	Y_i^H	Y_i^l	μ	δ
Type 1	0.5	130	20	75	110
Type 2	0.5	130	20	75	110
Type 3	0.5	130	20	75	110

$HE\delta$	p	Y_i^H	Y_i^l	μ	δ
Type 1	0.5	150	0	75	150
Type 2	0.5	130	20	75	110
Type 3	0.5	110	40	75	70

$HE\mu\delta$	p	Y_i^H	Y_i^l	μ	δ
Type 1	0.5	216	0	108	216
Type 2	0.5	130	20	75	110
Type 3	0.5	52	32	42	20

$HE\mu$	p	Y_i^H	Y_i^l	μ	δ
Type 1	0.8	130	20	108	110
Type 2	0.5	130	20	75	110
Type 3	0.2	130	20	42	110

In the benchmark HO treatment all participants face identical risks. Risk sharing thus implies that income is redistributed from the fortunate to the less fortunate individuals ex-post. In the $HE\delta$ treatment individuals also face the same μ, however, lottery spreads are different with $\delta_1 > \delta_2 > \delta_3$. Hence, this treatment examines simple mean-preserving spreads Rothschild and Stiglitz, 1970, a special case of second-order stochastic dominance. In the $HE\mu$ treatment lottery spreads stay constant but the expected lottery values increase across lottery types with $\mu_1 > \mu_2 > \mu_3$. Individuals face different probabilities for the same possible lottery outcomes and these probability shifts are first order stochastic dominance shifts. The $HE\mu\delta$ treatment combines the previous two forms of risk heterogeneity. In this treatment, as the spread increases, $\delta_1 > \delta_2 > \delta_3$, also the mean increases, $\mu_1 > \mu_2 > \mu_3$, across lottery types. Hence, there is a simple compensated risk increase, which does a priori not allow for a preference ranking of the lottery types. In order to make treatments better comparable, the overall expected income within a risk sharing group is the same for all treatments. Moreover, the expected lottery values are identical in treatments $HE\mu$ and $HE\mu\delta$ for each lottery type. Importantly, the risk exposure of lottery type two is identical across treatments.

3.2.2 Risk Sharing Mechanism

Individual lotteries are public information to all group members. Each participant is asked to choose a risk sharing level $s_i \in [0,1]$ which represents the fraction of his eventual income that is placed in the group account $G = \sum_{i=1}^{3}(s_i \cdot Y_i)$. This group account is then proportionally divided among all group members. An individual's earnings are consequently defined as

$$\Pi_i(s_i, s_{-i}) = (1 - s_i) \cdot Y_i + r_i \cdot G \tag{3.1}$$

with $r_i = \frac{s_i}{\sum_{j=1}^{3} s_j}$ being the percentage that an individual receives from the group account. The more risk is shared, the more incomes are equalized ex-post. If all individuals choose complete risk sharing, $s_i = 1$, eventual earnings of all group members will be identical. The total amount earned in a group, however, is uncertain, since risks are not correlated. If all individuals decide not to share risks at all, each participant's earnings will solely depend on his individual lottery outcome.

3.2.3 Procedures

Each participant was randomly assigned to a computer in the laboratory based on a random draw from a pile of numbered cards. Instructions were provided on the computer and a summary of the most important aspects was provided on paper.[1] Thereafter, participants were asked to answer a few comprehension questions concerning the group account, individual payments, returns and final incomes in order to ensure their understanding of the risk sharing mechanism. When all subjects had correctly answered the questions, the experiment was started. Participants were asked to make risk sharing decisions in three consecutive periods in the same treatment. The group constitution changed in each period as well as the assigned lottery type, so that each participant was assigned each lottery type once.[2] No feedback about lottery outcomes or sharing decisions was provided until the very end of the experiment.

In addition to the risk sharing decisions in each period we also elicited subjects' beliefs about the other group members' decisions by asking them to state the intervals they expect the sharing decision of the other two group members to be in. Belief elicitation was incentivised using the truncated interval scoring rule.[3]

If the actual sharing decision lay between the interval borders I_l and I_H, the participant was rewarded based on its length: the smaller the interval, the higher the payoff. If the actual decision was outside the interval the participant's payoff was zero. For our analysis we calculated an individual i's belief about individual j's risk sharing level as $\hat{s}_j = \frac{I_{jH}+I_{jl}}{2}$ $(i \neq j)$.

In the second part of the experiment we elicited individuals' risk preferences. We applied the multiple choice list method (Harrison and Cox, 2008). Certainty equivalents were elicited for the seven different lotteries that sub-

[1] See Appendix 3.6.2 for the instructions.

[2] The assignment of the lottery types was random in period one and thereafter changed such that each lottery type would be represented in each of the three periods. The participants being lottery type 1 in the first period were lottery type 2 in the second period and lottery type 3 in the third period. The orders for the other two lottery types were 2, 3, 1, and 3, 1, 2, respectively.

[3] It can be shown that with this scoring rule the specified interval contains the events subjects believe most likely to occur and that this property holds for risk neutral as well as risk averse subjects Schlag and van der Weele, 2013.

jects could face in the first part. For each lottery individuals were asked to make 20 choices between a sure amount and the respective lottery. The sure amount varies between Y_i^H and Y_i^l. Certainty equivalents are calculated as the arithmetic mean of the smallest sure amount preferred to the lottery and the consecutive sure amount on the list. Each individual's risk attitude α was then computed by minimizing the squared distances between the certainty equivalents observed and those theoretically predicted by expected utility theory assuming a utility function for money $U(x) = x^\alpha$ (see Wakker, 2008, Wakker, 2010). That is:

$$\min_{\alpha} \sum_{i=m}^{7} [(p_m(Y_m^H)^\alpha + (1-p_m)(Y_m^l)^\alpha)^{\frac{1}{\alpha}} - ce_m]^2$$

with ce_m being the elicited certainty equivalent of lottery $m = 1,..7$. The resulting measure is $\alpha_i = 1$ for risk neutral individuals, $\alpha_i < 1$ for risk averse individuals and $\alpha_i > 1$ for risk seeking individuals.

Lastly, participants were asked to answer a short questionnaire, were provided with information about lottery outcomes and earnings and were then paid in cash and dismissed from the laboratory. Participants were recruited online with the system ORSEE (Greiner, 2004). The computerized experiment was conducted in the Behavioural & Experimental Economics laboratory of Maastricht University (BEElab) using the experimental software z-Tree (Fischbacher, 2007). The experimental earnings were calculated as Experimental Currency Units (ECU) with a conversion rate of 1 ECU=0.04 Euros. In total, 120 students (30 in each treatment) participated in the one and a half hour long experiment. Average earnings were approximately 18 Euros.

3.3 Theoretical background

The following theoretical discussion is based on expected utility theory. In our experiment each of the three individuals $i = 1,2,3$ in a group faces a lottery which leads to a high income Y_i^H with probability p_i and to a low income Y_i^l with probability $(1-p_i)$. Each individual independently decides on how much risk to share, by deciding on the share s_i of his eventual (ex-post) income he puts into the group account. This share, together with the sum of

shares of the other individuals, determines the individual's return from the group account. To make an optimal sharing decision each individual has to consider the (expected) sharing decision of others.

More formally, let $\hat{s}_{-i}$ denote i's beliefs about the other group members' sharing decisions. When making the sharing decision, each individual has to consider the overall lottery with the eight possible realizations $\Pi_{ik}(s_i, \hat{s}_{-i})$ $(k = 1, 2, ..., 8)$ for the different combinations of high and low incomes together with the expected sharing decisions of others. Each of these possible realizations is given by

$$\Pi_{ik}(s_i, \hat{s}_{-i}) = (1 - s_i) \cdot Y_{ik} + \frac{s_i}{s_i + \sum \hat{s}_{-i}} \cdot (\sum \hat{s}_{-i} \cdot Y_{-ik} + s_i \cdot Y_{ik}), \tag{3.2}$$

where Y_{ik} denotes individual i's ex-post income (high or low) in the k^{th} realization. With p_k denoting the probability with which each of the eight realizations occur, an expected utility maximizing individual i will choose s_i as follows:

$$\max_{s_i} EU_i(s_i, \hat{s}_{-i}) = \sum_{k=1}^{8} p_k \cdot U[\Pi_{ik}(s_i, \hat{s}_{-i})]. \tag{3.3}$$

Risk sharing levels (s_1, s_2, s_3) constitute a Nash equilibrium if, for each i, s_i maximizes $EU_i(s_i, s_{-i})$ given s_{-i}. Note that $(s_1 = s_2 = s_3 = 0)$ is always a Nash equilibrium, independent of the treatment and independent of risk preferences. A player has no (strict) incentive to share risk if others do not share risk. To further analyze equilibrium sharing we assume that individuals' preferences can be represented by the power utility function $U(x) = x^\alpha$, and that preferences are symmetric and common knowledge. Under these assumptions it can be shown that in treatment *HO* full risk sharing $(s_1 = s_2 = s_3 = 1)$ is an equilibrium for all $\alpha < 1$. A risk averse player strictly prefers to fully share risk as soon as at least one other player shares (some) risk. In $HE\delta$ *joint* risk sharing $(s_1 > 0, s_2 > 0, s_3 > 0)$ is an equilibrium for all $\alpha < 1$, but full risk sharing is not. Lottery type 3, the one with the lowest risk, always prefers to share some risk but is never willing to fully share. In $HE\mu\delta$ joint risk sharing $(s_1 > 0, s_2 > 0, s_3 > 0)$ is an equilibrium only for some $\alpha < 1$. Lottery type 1, the one with the highest mean income as well as the highest income variance, is willing to share only if s/he is sufficiently risk averse. Finally, in $HE\mu$ joint risk sharing is never an equilibrium because lottery type 1, the one with the highest mean income, is unwilling to

share risk (i.e., $s_1 = 0$) even if s/he is very risk averse. Summarizing, in the aggregate we expect that the level of joint risk sharing is highest in HO, second highest in $HE\delta$, third highest in $HE\mu\delta$, and lowest in $HE\mu$.

3.4 Results

In this section we first report the aggregate risk sharing decisions and how they change across treatments and lottery types. Second, we compare the empirical results to theoretical benchmark conditions taking into account individual risk preferences and beliefs about others' sharing decisions. Finally, we explore potential strategic adverse selection by investigating how beliefs about other risk sharing decisions affect one's own sharing decision.

3.4.1 Risk sharing and adverse selection

At the moment of their risk sharing decisions subjects do not know the realizations of their lotteries. Hence, a natural measure of aggregate risk sharing is the ex-ante expected amount of money in the group account: the sum over individuals' risk sharing percentages times the expected value of their lotteries.[4] Using this measure, Figure 3.1 shows that risk sharing is highest in HO and $HE\delta$, smaller in $HE\mu\delta$ and smallest in $HE\mu$. This is roughly in line with the theoretically predicted ranking, except that risk sharing levels do not differ between HO and $HE\delta$. A Kruskal-Wallis (KW) test indicates that significant differences between treatments exist ($p = 0.0347$).[5] Pairwise comparisons with Mann-Whitney (MW) tests reveal that in $HE\mu$ the expected amount in the group account is significantly smaller than in HO and $HE\delta$ ($p \leq 0.0283$; FDR-corrected at 10% level). Other pair-wise compar-

[4] An alternative measure would be the expected amount of transfers implied by a risk sharing scheme. Using this measure leads to similar results but it is considerably more complex to calculate.

[5] Throughout the paper, whenever we carry out multiple pair-wise comparisons we report corrected significance levels using the false discovery rate (FDR), next to the uncorrected p-values. This method reduces the risk of false positives and controls for the rate of false negatives (Benjamini and Hochberg, 1995). We refrain from reporting FDR corrections when uncorrected p-values indicate insignificance. All reported tests are 2-sided.

isons (i.e., HO vs $HE\delta$, HO vs. $HE\mu\delta$, $HE\delta$ vs. $HE\mu\delta$, and $HE\mu\delta$ vs. $HE\mu$) yield insignificant differences when FDR-corrected.

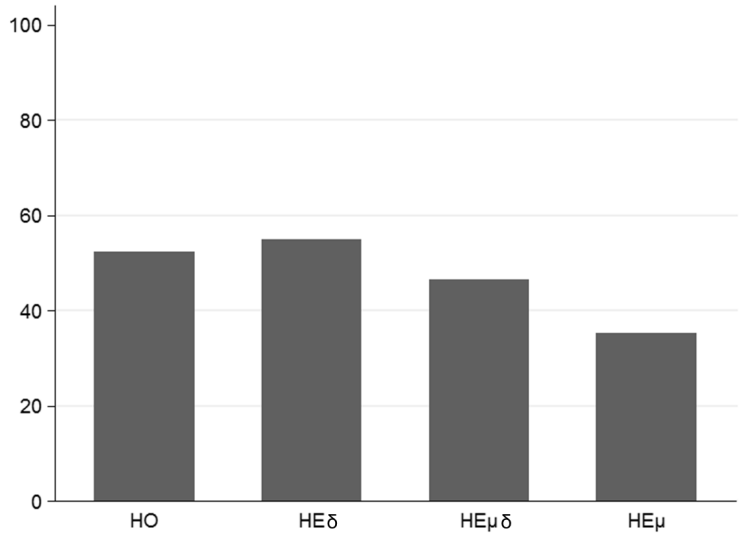


Fig. 3.1: Expected amount in group account (in percent)

Result. RISK SHARING AND RISKINESS
(1) Compared to the theoretical benchmark, too little risk is shared when risk is equally distributed.
(2) As predicted risk sharing decreases with differences in individuals' risk profiles. Specifically, it is smallest when individuals' risk profiles differ according to first-order stochastic dominance.

The aggregate sharing decisions conceal considerable differences between different lottery types within and across treatments. Figure 3.2 shows average risk sharing separately for each lottery type in each treatment. As expected, in HO, where all individuals in a group have the same risk profile risk sharing decisions do not depend on lottery types (averages are 51.7%, 53.3%, and 52.3%, for type 1, 2, and 3, respectively). Wilcoxon signed-rank (WS) tests indicate that the three different lottery types exhibit statistically indistinguishable risk sharing decisions ($p \geq 0.992$).[6]

[6]Since we use the strategy method, we have three risk sharing decisions for each subject:

In $HE\delta$, where individual risk profiles can be ranked according to mean-preserving spreads, lottery types decrease in their riskiness from left (lottery type 1) to right (lottery type 3). Adverse selection would predict decreasing risk sharing in the same direction. Figure 3.2 shows that empirically risk sharing is highest when participants are assigned lottery type 2 (62.2%) followed by lottery type 1 (55%) and lottery type 3 (47.7%). A WS test yields that the difference in risk sharing between lottery type 2 and 3 is significant at the FDR-corrected 5% level ($p = 0.013$). As all other pair-wise comparisons return insignificant results ($p \geq 0.432$), we consider this as only weak evidence in favor of adverse selection when individuals' risk profiles can be ranked according to mean-preserving spreads.

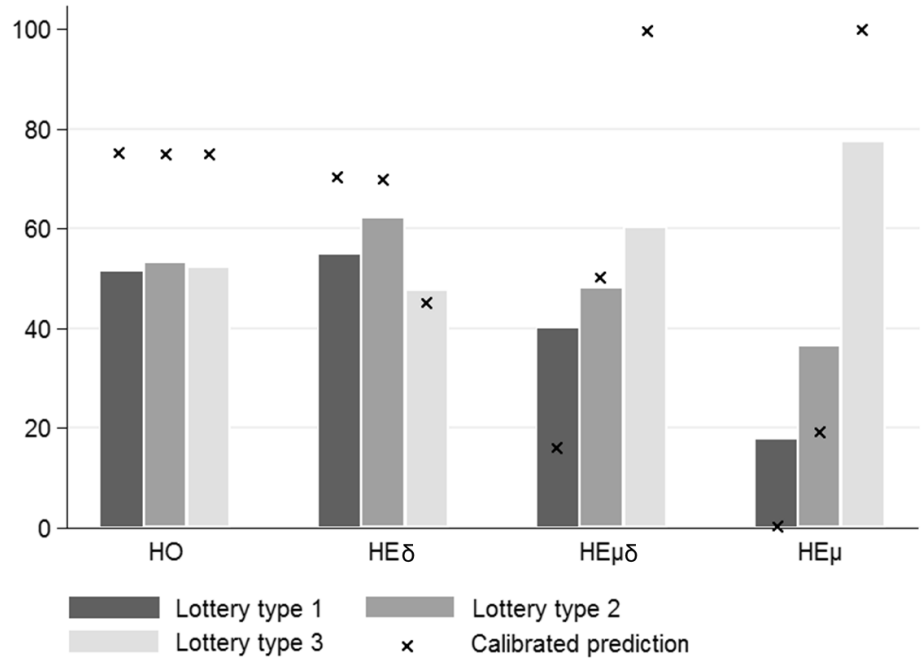


Fig. 3.2: Risk sharing per lottery type (in percent)

one for each lottery type. For each treatment, we use these 30 triplets of observations to perform three Wilcoxon matched-pair signed-rank tests to compare within-subject sharing decisions between lottery type 1 and lottery type 2, between type 1 and type 3, and between type 2 and type 3, respectively.

In $HE\mu\delta$ and $HE\mu$ results are very different. In both treatments, risk sharing levels of lottery types clearly increase from left (lottery type 1) to right (lottery type 3). This correlates negatively with the expected values of the lottery types, which increase from right to left, and clearly indicates the existence of adverse selection when risk profiles can be ranked according to first-order stochastic dominance. Specifically, in $HE\mu$ individuals with the low expected value lottery type 3 (bad risk) choose on average a risk sharing level of 77.4%, those with an intermediate risk profile (lottery type 2) choose a level of 36.5%, and those with a high expected value (good risk) choose a sharing level of only 17.9%. Pairwise WS tests reveal that differences between lottery types are statistically significant at the FDR-corrected 5% level ($p \leq 0.001$). With a difference of about 60 percentage points between the low and high expected value lottery type, adverse selection effects are also of economic significance.

In $HE\mu\delta$ risk increases from left (lottery type 1) to right (lottery type 3) in terms of expected values but decreases in terms of outcome spreads.[7] Hence, individuals face a simple compensated risk increase with opposing incentives from means and spreads. These countervailing effects on adverse selection are empirically reflected by less extreme differences in sharing across lottery types than in $HE\mu$. Still sharing is highest for lottery type 3 (60.3%), intermediate for lottery type 2 (48.1%), and lowest for lottery type 1 (40.1%). However, the difference between the two extreme lottery types is with 20.2 percentage points less stark than in $HE\mu$ and not all differences are statistically significant. Pairwise comparisons with WS tests show that between lottery type 3 and lottery types 1 and 2 the differences are statistically significant at the FDR-corrected 5% level ($p \leq 0.027$) but not between lottery type 1 and 2 ($p = 0.216$).

Result. ADVERSE SELECTION AND RISK PROFILES
(1) If individuals' risk profiles are ranked according to mean-preserving spreads little adverse selection is observed.
(2) If individuals' risk profiles are ranked according to first-order stochastic dominance there is strong evidence for adverse selection.

[7] In the following we will refer to the risk of lottery type 1 as the 'good' risk in $HE\mu\delta$ and $HE\mu$, whereas in $HE\delta$ lottery type 3 is the 'good' risk.

(3) The latter adverse selection effect is mitigated if individuals face a simple compensated risk increase (i.e., increasing mean and spread).

To get a better idea how the empirical results compare to theoretically predicted behavior we calculated the individually optimal risk sharing levels for each subject for each lottery type using the elicited individual risk preference and beliefs about others' sharing decisions.[8] Specifically, we assume that the preferences of each subject i can be represented by a utility function for money $U_i(x) = x^{\alpha_i}$, where the individual risk parameter α_i is estimated as described in Section 3.2.3. Further, we use the elicited beliefs of i about others' sharing decisions $\hat{s_j}$ to solve for each i the maximization problem in Equation (3.3), given the possible realizations in Equation (3.2).

In Figure 3.2 the black crosses indicate the average of these individually optimal risk sharing levels.[9] Two regularities are immediately eye-catching. First, for each treatment, across types the changes in sharing levels fit the theoretical comparative statics results qualitatively very well. 'Bad' risk lottery types are predicted to share more than intermediate risk lottery types, and in turn, intermediate lottery types are predicted to share more than 'good' risk lottery types. Second, quantitatively the empirical sharing levels sometimes diverge significantly from the individually optimal ones.

In HO actual risk sharing is significantly lower than theoretically predicted ($p \leq 0.032$ for all lottery types, WS tests; FDR-corrected at 5% level for lottery types 3 and 2, at 10% level for lottery type 1). In $HE\delta$, however, empirical risk sharing levels do not differ from predicted ones for any lottery type ($p \geq 0.128$, WS tests). In $HE\mu$ and $HE\mu\delta$ predicted adverse selection is stronger than is actually the case. In $HE\mu$, for all three lottery

[8] Note that this does not give an equilibrium prediction but rather a 'naïve' optimal decision for each individual. We consider such naïve behavior of our subjects as more realistic than equilibrium behavior for at least two reasons. First, the latter would ask for complex computations our subjects unlikely went through. Second, it also requires common knowledge of all individuals' risk preferences, information our subjects did not have and also could not acquire as all sharing decisions were only made once.

[9] When eliciting participants' risk preferences, 9 of the 120 individuals switched more than once between the lottery and the safe payment in at least one of the decision situations. We thus do not have a risk preference measure for them. The bars in Figure 3.2 are therefore based on the decisions of all 120 individuals, whereas the calibrated predictions and the according analysis are based on 111 observations only.

types differences between theoretical and empirical sharing levels are statistically significant at the FDR-corrected 5% level ($p \leq 0.024$, WS tests). In $HE\mu\delta$ differences between theoretically predicted and empirically observed sharing levels are significant for lottery types 1 and 3 ($p \leq 0.001$, WS tests; FDR-corrected at 5% level) but not for lottery type 2 ($p = 0.715$, WS test).

Result. INDIVIDUALLY OPTIMAL AND EMPIRICAL ADVERSE SELECTION
(1) Qualitatively, empirically observed adverse selection coincides with individually optimal adverse selection: risk sharing correlates with risk profiles in the predicted way.
(2) Quantitatively, adverse selection is less pronounced than individually optimal risk sharing would dictate.

3.4.2 Anticipated and strategic adverse selection

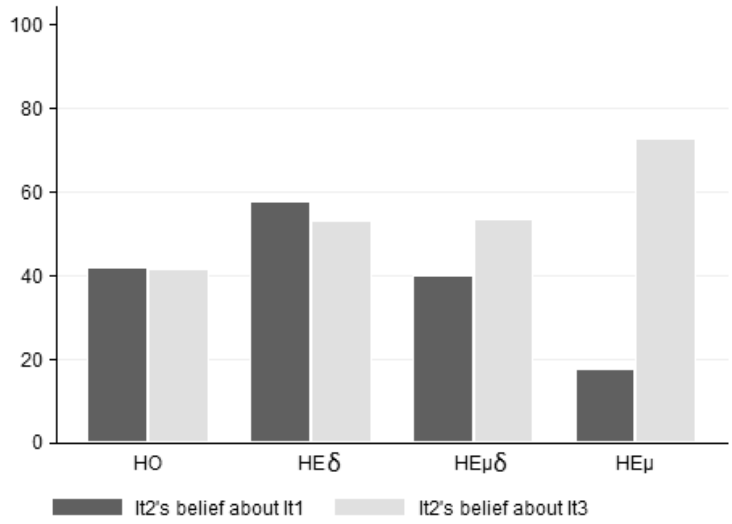
An important element of our experimental design is that individuals with an intermediate risk profile (lottery type 2) faced exactly the same individual lottery and, hence, risk exposure in all treatments, whereas the risk profiles of the other lottery types systematically varied across treatments. Moreover, the sum of the expected values of lottery types 1 and 2 is also identical in all treatments.

In Section 3.4.1 we have seen that risk sharing of individuals with intermediate risk profiles is highest in HO and $H\delta$, lowest in $H\mu$, and intermediate in $H\mu\delta$.[10] This strongly suggests that these differences are a response to anticipated adverse selection and, hence, strategic. We use the beliefs that subjects with intermediate risk profiles have about other lottery types' risk sharing to investigate whether adverse selection is indeed anticipated. If so, we should see differences in beliefs of lottery type 2 regarding the other types' risk sharing levels across treatments, which should correlate with actual risk sharing of lottery types 1 and 3.

Figure 3.3 reports the average risk sharing levels of lottery types 1 and 3, expected by lottery type 2. It shows that beliefs about other types are clearly in line with these types' actual behavior (cf. Figure 3.2). There are no statistically significant differences between type 2's beliefs and the actual

[10] The differences in sharing levels are FDR-corrected (marginally) significantly different between $H\mu$ and HO and $H\delta$ and between $H\delta$ and $H\mu\delta$ ($p \leq 0.0318$, MW tests).

behavior of types 1 and 3 ($p \geq 0.1985$, MW tests). This is corroborated by a regression analysis showing that risk sharing of lottery type 2 individuals is positively correlated with the anticipated risk sharing by the 'good' risk lottery type (see Appendix 3.6.1).



Note: The labels "lt$X$'s belief about lt$Y$" should be read as "subject with lottery type X believes that risk sharing of subject with lottery type Y is ...".

Fig. 3.3: Belief of lottery type 2 about other lottery types' sharing levels (in percent)

Result. RISK PROFILES, ANTICIPATED AND STRATEGIC ADVERSE SELECTION
(1) Intermediate risk types, with a constant risk profile, exhibit varying degrees of adverse selection across treatments.
(2) Intermediate risk types exhibit strategic adverse selection. That is, their own risk sharing correlates with their correctly anticipated adverse selection behavior of other risk types.

3.5 Conclusion and Discussion

We experimentally investigate how heterogeneity in risks affects voluntary risk sharing. Participants are asked to decide how much risk they are will-

ing to share, knowing their and others' risk profiles but facing uncertainty about risk outcomes. The results provide behavioral evidence for the relevance of adverse selection. Individuals with a 'good' risk profile are less likely to enter a risk sharing arrangement than individuals with a 'bad' risk profile. The strength of the effect, however, depends on the manner in which the risk profiles vary between individuals. Risks which differ in terms of mean outcomes (first-order stochastic dominance) generate much stronger selection effects than risks which differ by way of mean-preserving spreads.

Qualitatively, our results are in line with the theoretical predictions, assuming CRRA utility maximization. The observed treatment effects line up nicely with the theoretical comparative statics. Risk sharing levels are higher in treatments in which risk sharing is an equilibrium under a wider set of parameters. Moreover, calibrated predictions at the individual level using elicited risk preferences and beliefs correlate strongly with observed decisions. Quantitatively, however, point predictions are often off the mark. Specifically, we find that risk sharing levels are generally less extreme than predicted. For example, in a symmetric environment with homogenous risks, risk sharing levels are about 30%-points lower than predicted. Also, selection effects generally fall short of the prediction.

It should be noted that our experiment is single shot, and that participants have no opportunity for learning. Hence, some subjects may resort to the focal contribution level of 50 percent. If that is indeed the case, our results on adverse selection are likely to be a lower bound as gravitation towards the midpoint contribution mitigates adverse selection effects. It would be interesting to investigate this in future work and explore how risk sharing and adverse selection develop over time with feedback and the possibility of learning.

Another potential explanation for the difference between the observed risk sharing levels and the calibrated predictions could be that (some) subjects have social preferences (e.g. Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). In the homogeneous risk case too little risk-sharing could be consistent with disadvantageous inequality aversion when one expects others to share little risk. We indeed find that subjects expect others to share relatively little risk. On the other hand, in the heterogeneous risk treatments, we also see too little adverse selection of low risk subjects which may

also point at advantageous inequality aversion. As our experiment was not designed to disentangle these potential explanations, future research will have to show which of these or other motivations best explain our results.

In many countries there are continuing debates on whether mandatory insurance should be introduced for some risks (think of the Affordable Care Act in the United States), while in other countries (partial) exit options are being proposed for certain collective arrangements (such as mandatory occupational pension schemes). A key issue in the public debate, at least in Europe, is whether the option to opt out may undermine 'solidarity', where the lucky support the unlucky. Although one must be cautious to generalize from lab to field, our results suggest that such concerns are well-founded behaviorally. At the same time, our results indicate that much will depend on the manner and degree to which risk profiles differ. The easier it is to order different risk profiles from 'good' to 'bad', the stronger selection effects are likely to be.

Finally, one may question whether our design is somehow biased towards finding certain effects. Indeed, in our experiment participants' risk profiles are common knowledge. This may lead to an overestimation of the importance of selection for environments in which people do not have very precise knowledge about how their risk profile compares to that of others. On the other hand, however, in our experiment risk profiles are assigned exogenously. This may underestimate the force of adverse selection for settings in which risks are endogenous. The support for redistribution is typically higher in case poverty is caused by bad luck rather than a lack of effort (Cappelen et al., 2013). Similarly, low risk individuals may be (even) less willing to share risk with high risk individuals in case they perceive that those individuals are accountable for their risk (see Cettolin and Tausch, 2013). Future work will have to establish which of these elements are more or less important for the behavioral relevance of adverse selection.

3.6 Appendix

3.6.1 Additional regression analysis

We observe that the more individuals with lottery type 2 expect the 'good' risk lottery type to share, the more they share themselves. This can be seen from the regression results below, where we only use observations for lottery type 2. The dependent variable is lottery type 2's risk sharing level. The variable Belief_goodrisk is lottery type 2's belief about the sharing level of the good risk type.

Table 3.1: Effect of anticipated risk sharing of 'good' risk type on lottery type 2's sharing decision

Variable	**Coefficient**	**(Std. Err.)**
$HE\mu\delta$	-8.826	(5.596)
$HE\mu$	-9.815	(6.470)
Gender	-4.220	(4.482)
Age	1.006	(1.093)
Belief_goodrisk	0.488**	(0.097)
Intercept	32.496	(24.612)
N	90	
Adj. R^2	0.36	

Note: ** 5% significance level; treatment HO not included; treatment $HE\delta$ is the omitted treatment category; session dummies included but not reported.

3.6.2 Experimental Instructions

The following instructions appeared on participants' computer screens (example for treatment $HE\mu$).

Risk Sharing Part

The decision task you are going to face will proceed as follows: You are randomly assigned to a group consisting of three participants including yourself.

Each of you will be assigned the role of participant A, B or C and face one of the following three lotteries:

With 50% chance participant A receives 130 ECU and with 50% chance participant A receives 20 ECU.

With 20% chance participant B receives 130 ECU and with 80% chance participant B receives 20 ECU.

With 80% chance participant C receives 130 ECU and with 20% chance participant C receives 20 ECU.

Each lottery outcome will be randomly and independently determined. That is, your lottery outcome does in no way depend on any of the other lottery outcomes. If your probability of getting the high outcome is 80%, it means that in 80 out of 100 cases you will get the high outcome. If it is 20%, it means that in 20 out of 100 cases you will get the high outcome and if it is 50%, it means that in 50 out of 100 cases you will get the high outcome.

1. You are informed about the lottery that has been assigned to you and the lotteries assigned to the other two group members. However, you do not know the lottery outcomes.
2. You are asked to decide which percentage of your lottery outcome you want to put into a group account. This decision is called X. You can choose any percentage between 0% and 100% in steps of one percent: X = 0%, 1%, 2%, 3%, ...,98%, 99%, or 100%. The amount in ECU you put into the group account is calculated as the percentage X times your lottery outcome. The rest, (100-X)% times your lottery outcome, will be added to your private account.
3. In the group account there is the sum of what all three group members put into it. The proportion a member gets back from the group account is zero whenever he/she puts 0% into the group account. Otherwise, the group account is proportionally divided among the three group members. That is, the percentage of the group account a member gets back from the group account is calculated as the member's own choice X divided by the sum over all three choices X. The higher your choice X, the higher the proportion of the group account you get back. For example, if you choose X= 0%, you put your individual lottery outcome in your private account and you do not add

anything to the group account. In this case you do not get anything back from the group account. Your X is 0 and what you get back is X/(sum of all three X) = 0/(sum of all three X)= 0. If you choose X= 100%, you put your whole lottery outcome into the group account. In this case what you get back is 100/(sum of all three X). Thus, you choose the highest share of the group account that is possible.
4. Your final outcome for this part of the experiment will be calculated as the sum of your private account plus your return from the group account. Note, that the total amount in the group account also depends on the decisions of the other three members!

- Comprehension questions are answered -

[Instructions Period 1]

On the following screens you will be informed about the lottery that is randomly assigned to you as well as the lotteries that are randomly assigned to the other two group members.

You will be asked what percentage of your lottery outcome you put into the group account. When making this decision your lottery outcome and the other participant's lottery outcomes have not yet been randomly determined. The other two group members are faced with the same decision situations. The earnings in a decision situation will depend on all group members decisions in the way just described in the practice examples.

You will be asked to report your best guesses of how each of the other two members will decide. Specifically, for each other member you will be asked to report the interval in which you believe their decisions will lie. That is, you will report your best guess of the smallest and largest percentage each of the other two will put into the group account. With these guesses you can earn extra money and you will earn the more money the more accurate your guess is.
Your earnings from this task are determined in the following way: If another's actual chosen percentage is outside your guessed interval (that is, is smaller than your smallest guess or larger than your largest guess) then you earn nothing with your guess. If another's actual chosen percentage is inside your guessed interval (that is, is equal or larger than your smallest

guess and equal or smaller than your largest guess) then you earn the more the smaller your chosen interval is. For instance, if you guess one particular percentage $Z\%$ (the interval $Z\% - Z\%$), and the other member indeed puts this particular percentage $Z\%$ into the group account then you earn the maximal amount of 15 ECU with your guess. If you guess that the percentage will lie in the interval $0\% - 100\%$ then you will earn 0 ECU. For intermediate intervals you earn in proportion to the length of the interval. For example, if you choose $28\% - 48\%$ and the other's chosen actual percentage is 32% (that is, inside the interval) then you earn $15 - 15 * ((48 - 28)/100) = 15 - 15 * 0.2 = 12$ ECU; if you would have chosen $16\% - 67\%$ you would have earned $15 - 15 * 0.51 = 7.35$ ECU

The total earnings are the sum of your earnings resulting from your and the others actual decisions and the earnings resulting from your guesses.

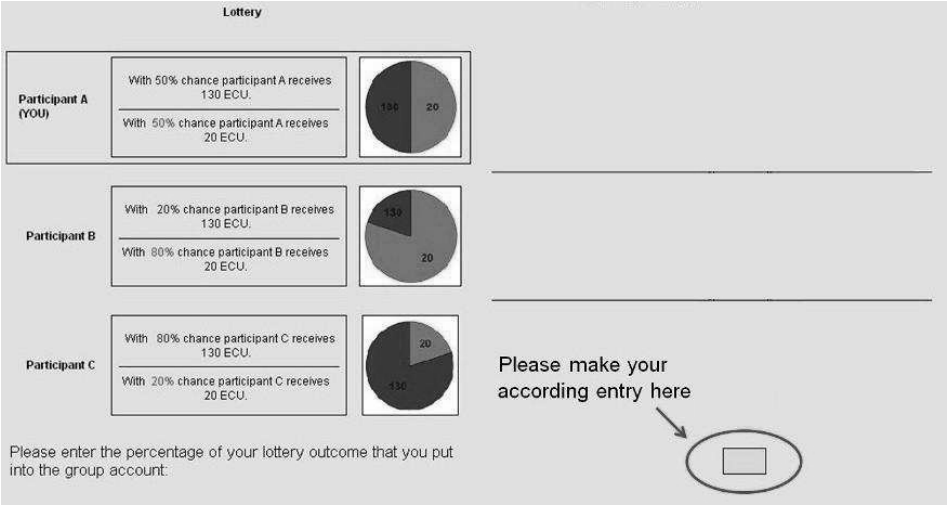


Fig. 3.4: Screenshot first decision task

This is an example screen shot of the first decision task. You are not asked to make choices now! Please have a careful look. In this example you are participant A. You are asked to enter in the field below, emphasized by the red round frame, what percentage of your lottery outcome you want to put into the group account.

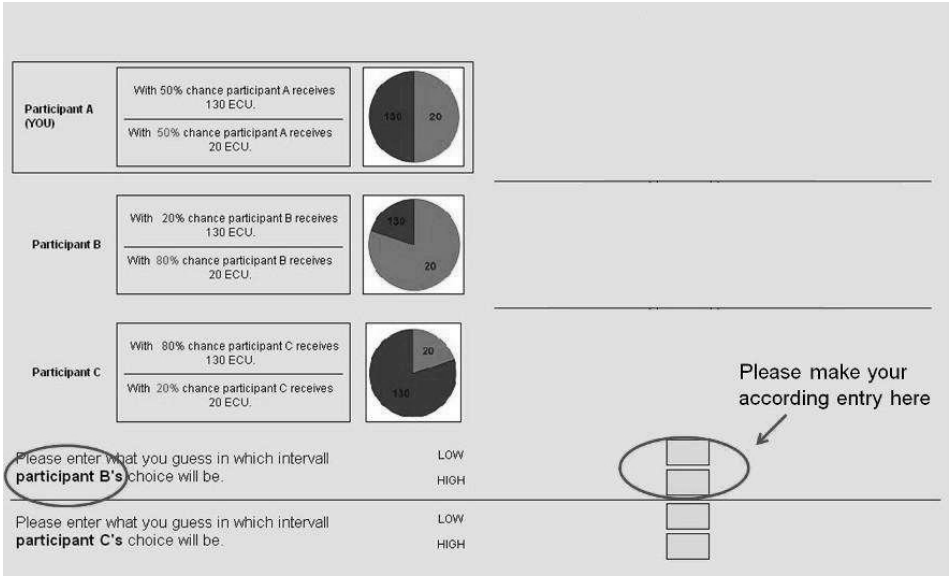


Fig. 3.5: Screenshot second decision task

This is an example screen shot of the second decision task. You are not asked to make choices now! Please have a careful look. In this example you are asked to enter in the field emphasized with the red round frame, what you think in which interval participant B's choice X will be in that particular decision situation. In the field next to 'LOW' you enter the lower bound of the interval, in the field next to 'UP' you enter the upper bound of the interval. Take care that you don't exchange them! You do the same for the guess about participant C's decision below.

You now enter the actual decision stage! The entries you make on the next screens will be relevant for your payment.

- Decisions for period 1 are made -

[Instructions Period 2]

You are now again matched with two other participants and confronted with the same decision situation as on the previous screens. However, this time you face a different lottery than in the first part. Notice that you will not

be with the same participants in one group again! The entries you make on the next screens will be relevant for your payment as well. The resulting earnings will be added to your earnings from the first part.

- Decisions for period 2 are made -

[Instructions Period 3]

You are now again matched with two other participants and confronted with the same decision situation as on the previous screens. However, you face a different lottery than in the first two parts. Notice that again you will not be with the same participants in one group twice! The entries you make on the next screens will be relevant for your payment as well. The resulting earning will be added up to your earnings from the first and second part.

- Decisions for period 3 are made -

Risk Elicitation Part

You are now going to make another series of choices. These choices will not influence your earnings from the choices you just made, nor will your earlier choices influence the earnings from the choices you are going to make. After you have made these choices you will be asked to answer some questions. Thereafter the experiment will be over. In the following, you will be confronted with a series of 7 decision situations that will appear in random order on the screen. All these decision situations are completely independent of each other. A choice you made in one decision situation does not affect any of the other following decision situations.
Each decision situation is displayed on a screen. The screen consists of 20 rows. You have to decide for every row whether you prefer option A or option B. Option A is a lottery and is the same for every choice in a given decision situation, while the secure option B takes 20 different values, one for each choice.

This is a screen shot of one arbitrarily chosen decision situation you are going to face. You are not asked to make choices now! Please have a careful look. At the end of the experiment one of the 7 decision situations will be

DECISION SITUATION 5.	OPTION A LOTTERY	YOUR CHOICE	OPTION B SURE AMOUNT
choice 1	With 20% chance you receive 130 ECU, with 80% chance you receive 20 ECU. 	A ○ ○ B	130 ECU
choice 2		A ○ ○ B	125 ECU
choice 3		A ○ ○ B	120 ECU
choice 4		A ○ ○ B	115 ECU
choice 5		A ○ ○ B	110 ECU
choice 6		A ○ ○ B	105 ECU
choice 7		A ○ ○ B	100 ECU
choice 8		A ○ ○ B	95 ECU
choice 9		A ○ ○ B	90 ECU
choice 10		A ○ ○ B	85 ECU
choice 11		A ○ ○ B	80 ECU
choice 12		A ○ ○ B	75 ECU
choice 13		A ○ ○ B	70 ECU
choice 14		A ○ ○ B	65 ECU
choice 15		A ○ ○ B	60 ECU
choice 16		A ○ ○ B	55 ECU
choice 17		A ○ ○ B	50 ECU
choice 18		A ○ ○ B	45 ECU
choice 19		A ○ ○ B	40 ECU
choice 20		A ○ ○ B	35 ECU

Fig. 3.6: Screenshot risk elicitation

randomly selected with equal probability. Once the decision situation is selected, one of the 20 rows in this decision situation will be randomly selected with equal probability. The choice you have made in this specific row will determine your earnings. Consider, for instance, the screen shot that you have just seen. Option A gives you a 20% chance to earn 130 ECU and a 80% chance to earn 20 ECU. Option B is always a sure amount, in this case ranging from 130 ECU in the first row, to 35 ECU in the 20th row. Suppose that the 12th row is randomly selected. If you would have selected option B, you would receive 75 ECU. If, instead, you would have selected option A, the outcome of the lottery determines your earnings. Please note that each decision situation has the same likelihood to be the one that is relevant for your earnings. Therefore, you should view each decision independently and consider all your choices carefully.

- Decisions for risk elicitation are made -

Chapter 4

Risk Taking and Risk Sharing: Does Responsibility Matter?*

*This chapter is joint work with Elena Cettolin and currently under review. We thank Arno Riedl, Jan Potters, Ronald Peeters and Matthew Embrey for their helpful comments. We also thank the participants of the 7th Nordic Conference on Behavioral and Experimental Economics in Bergen, the 3rd Maastricht Behavioral and Experimental Symposium, seminar attendants at LMU Munich and at University of Tilburg for their suggestions.

4.1 Introduction

The fundamental premise for the support of safety nets, such as social security systems and private insurances, is that individuals are willing to share risk with others, thereby accepting the resulting redistribution of income. Indeed, whenever risk is shared those who are lucky support the more unlucky individuals in society.

For a long time, the idea of tailoring insurance rates to risk types has been debated in public.[1] For example, proposals to charge higher health insurance premiums to smokers and obese people have been advanced, with the motivation that a high proportion of health care costs can be directly attributed to patients' bad habits (see Cawley and Ruhm, 2011 and Thomson Reuters, 2011). In light of this evidence, we suggests that support for risk sharing arrangements is weak when individuals are perceived to be responsible for their risk exposure.

The decision to share risk may be backed by both insurance and redistribution motives. The first has a selfish nature, as it allows risk averse individuals to reduce their risk exposure. The second is driven by a preference for equality, as the more risk is shared the more income inequalities are reduced ex-post. We hypothesize that in the absence of responsibility attributions for risk exposure, redistribution motives are stronger and the willingness to share risk higher, as compared to when individuals can influence the risk they face. We test this conjecture using a controlled laboratory experiment, focusing on endogenous and exogenous differences in risk exposure. Our set up allows studying how the support for risk sharing depends on individuals' risk preferences, their own risk exposure, as well as their sharing partner's risk exposure.

Empirical research on risk sharing has identified a number of factors that affect individuals' propensity to share risk, e.g. group size (Chaudhuri et al., 2010), group selection and commitment (Barr and Genicot, 2008), risk preferences and social networks (Attanasio et al., 2012), one's own and others' risk profiles (Tausch et al., 2013) and reciprocity in repeated interactions (Charness and Genicot, 2009). However, to the best of our knowledge, we are

[1] For recent articles see New York Times on the web (2011), CNN on the web (2011), The Washington Post (2012).

the first to investigate how risk sharing depends on whether individuals perceive themselves and others to be responsible for the extent to which they are exposed to risk. Our results help to understand whether perceived choice responsibility is a crucial variable influencing the support of modern safety nets.

The experiment consists of two treatments. In the Exogenous Risks (EXO) treatment subjects cannot influence the extent to which they are exposed to risk, while in the Endogenous Risks (ENDO) treatment subjects can choose their risk exposure. In the first part of the ENDO (EXO) treatment subjects choose (are assigned) one of two risky lotteries. Both lottery options have the same expected value but differ in their variance. In the second part of both treatments, subjects are paired and one subject in each pair is randomly selected to choose a risk sharing level. Importantly, the risk sharing decision is made ex-ante, that is before the lotteries' outcomes are determined. The risk sharing level indicates the percentage amount that will be subtracted from the eventual outcomes and then equally redistributed in the pair at the end of the experiment. We implement the strategy method, which means that participants are asked to choose a risk sharing level both for the case that their partner faces the same risk exposure as themselves, and for the case that risk exposures differ. In the last part of the experiment we use a series of incentivised lottery choices to elicit participants' risk preferences.

Our main result is that in ENDO individuals systematically condition their risk sharing decisions on the risk exposure chosen by their partner, while the risk sharing behavior of individuals in EXO does not depend on the partner's risk exposure. Further, we find that average risk sharing is not significantly different in EXO than in ENDO.

Our research is related to some experimental studies that investigate the support for ex-post income redistribution in contexts where individuals' outcomes are the product of risky decisions. In Cappelen et al. (2013) participants make choices between a risky lottery and a safe alternative and after observing the eventual outcomes, they are asked how much they want to redistribute to another randomly matched participant. The authors find that individuals who avoid risk do not redistribute much in favor of unlucky risk takers, while the willingness to reduce inequalities is higher between lucky and unlucky risk takers. Thral and Rademacher (2009) implement the sol-

idarity game of Selten and Ockenfels (1998) and compare it to a treatment where individuals choose between a safe payment and a risky lottery. The authors show that individuals that choose the safe payment are less willing to reduce inequalities when matched with individuals that choose the lottery and become needy, as compared to individuals that become needy by pure chance. To summarize, it seems that risk taking is negatively perceived by individuals that avoid risk, and thus reduces their willingness to equalize earnings ex-post. Importantly, in the cited literature, redistribution decisions are made at a point when risk is resolved and individuals' outcomes are thus known. Our experiment allows testing whether responsibility for risky choices also matters when individuals face uncertainty about how risks eventually materialize. We believe that such an ex-ante perspective is worth investigating because most life outcomes are uncertain and hence, individuals often need to decide whether to support a given redistributing system before they observe their own outcome.

Our paper is also related to recent studies showing that income inequalities are more acceptable when they can be traced back to factors within peoples' control. Surveys, as well as experiments, reveal that support for redistribution is higher among people that think that wealth results from unjust motives, like luck or immoral behavior, as opposed to hard work, effort and skills (Alesina and Glaeser, 2004, Alesina and Ferrara, 2005, Fong, 2011, Durante and Putterman, 2009, Krawczyk, 2010). These results lend additional support for the idea that willingness to share risk may be related to whether risk exposure is perceived to be an exogenous factor or, on the contrary, an individual choice variable.

The remainder of the paper is organized as follows. Section 4.2 describes the experimental design. Section 4.3 summarizes theoretical predictions and hypotheses. Results are presented in Section 4.4. In Section 4.5 we discuss the results and conclude.

4.2 Design and procedures

We implement two treatments, EXO and ENDO, that differ with respect to whether subjects can choose the extent to which they are exposed to risk. Both treatments consist of three parts. In the following we describe each

part in detail and point out the treatment differences.

In the **risk exposure** part all subjects in ENDO make a choice between two lottery options, while in EXO subjects are assigned one of the two lotteries by a random draw operated by the computer. Subjects face a lottery choice (ENDO) or a lottery assignment (EXO) in four situations, that differ in the available lotteries.[2] We employ more than one situation in order to test whether results are robust to different combinations of outcomes and probabilities. In each of the four situations participants are presented with two lotteries, R (high risk) and r (low risk).[3] Both lotteries yield a high outcome, H, with probability p and a low outcome, l, with probability $1-p$. All the employed lotteries have the same expected value of €6, but in each situation the variance of lottery r is lower than that of lottery R.[4] In other words, lottery r second order stochastically dominates lottery R, and it is thus preferred by individuals with risk averse preferences. Keeping all lotteries' expected values equal ensures that situations differ only in one dimension, namely the difference in variance between lottery options. Hence, eventual differences in risk sharing across decision situations will be easier to interpret. Further, by abstracting from differences in expected values we obtain rather conservative results: in a given situation responsibility for risk taking is limited to the lottery's riskiness, as in expectation the contribution to the risk sharing pool is independent of the risk exposure. Table 4.1 gives an overview of the lotteries employed in the different situations. In each situation, probabilities and outcomes are selected in a way that participants can easily compare the two lotteries. In particular, lottery R and r are always equal in one dimension, either with respect to the outcomes' probabilities or with respect to the value of the lower outcome of the lottery, l.

In situations I and II, the riskier option implies either a higher probability of ending up with a zero outcome or introduces the possibility of a zero outcome, as compared to the safer alternative. Situations III and IV are less

[2] We also implemented a fifth situation where subjects choose between a safe payment and a risky lottery. The data referring to this situation are however not informative for our research question and are thus not included in the analysis of this paper.

[3] In the experiment a neutral wording is used. Please refer to the Appendix for the instructions used in the experiment.

[4] In situation II the expected value of r is €5.9. This exception was made to avoid confronting subjects with lottery outcomes that have more than one decimal point.

Table 4.1: Situations in the risk exposure part

situation	option	p	H	l
I	R	0.2	30	0
	r	0.5	12	0
II	R	0.6	10	0
	r	0.6	6.5	5
III	R	0.2	22	2
	r	0.2	10	5
IV	R	0.2	14	4
	r	0.5	8	4

extreme, in the sense that the low outcome of R is strictly larger than zero. We can thus test to what extent responsibility for risk exposure depends on the severity of the potential consequences of risk taking.

Situations are presented in random order to the participants. All participants are informed that only one of the five situations will matter for their final earnings. The instructions explain that each situation is equally likely to be selected for payment at the end of the experiment.

At the beginning of the **risk sharing** part subjects are randomly matched in pairs and in each pair one subject is selected at random to choose a risk sharing level $s \in [0,100]$. The value of s has to be chosen *ex-ante* and represents the percentage amount that is deducted from the lottery outcome of each subject, after risk is resolved. In each risk sharing pair, the deducted amounts are added up and equally re-distributed at the end of the experiment. Higher values of s imply lower levels of earnings' inequality, with $s = 100$ leading to equal outcomes in a pair. The following expression defines the earnings of a subject i resulting from the first two parts of the experiment, where Y_i is the lottery outcome of i and Y_j is the lottery outcome of i's risk sharing partner, j. The lottery outcomes of i and j are uncorrelated.

$$\Pi_i = (1 - \frac{s}{100}) \cdot Y_i + \frac{\frac{s}{100} \cdot (Y_i + Y_j)}{2} \tag{4.1}$$

We employ the strategy method to elicit risk sharing levels. That is, for each of the four situations in the risk exposure part, a subject chooses two

values of s. One for the case that the risk sharing partner faces the same lottery and one for the case that he faces the alternative lottery. Within a treatment, this allows observing whether individuals' risk sharing decisions are conditioned on the risk exposure of the risk sharing partner.

At the end of the experiment subjects are informed about the lottery chosen by (ENDO) or assigned to (EXO) their risk sharing partner. Further, the relevant risk sharing level is revealed to the subject who did not make decisions in the second part. Risk is then resolved, the chosen redistribution is implemented and earnings are determined. Instructions for the risk exposure and the risk sharing parts are administered together at the beginning of the experiment. Hence, in both treatments subjects know that decisions about risk sharing will have to be made after the risk exposure part.

Elicitation of Risk Preferences This part of the experiment is designed to estimate subjects' risk preferences. We use the multiple choice list method (Harrison and Cox, 2008) and elicit participants' certainty equivalents for the eight lotteries in Table 4.1. For each lottery subjects see a screen on the computer that contains a description of the lottery and a list of 20 equally spaced sure amounts, ranging from the lottery's high to its low potential outcome. In each row of the list subjects have to make a choice between the lottery and the sure amount. To ensure a unique switching point subject are not allowed to switch back and forth between the two. Certainty equivalents are then calculated as the arithmetic mean of the smallest sure amount preferred to the lottery and the consecutive sure amount in the list.

Experimental Procedures The experiment was conducted in the Behavioral and Experimental Economics Laboratory (BEElab) at Maastricht University. Subjects were recruited on line with the system ORSEE (Greiner, 2004). For the computerized implementation we used the experimental software Z-tree (Fischbacher, 2007). A typical session lasted approximately 1.5 hours and the average earnings were 18.70 Euro. In total 208 subjects participated in the experiment, 112 in the EXO and 96 in the ENDO treatment. In order to increase participants' understanding of the instructions a set of control questions was administered before the actual start of the experiment. Before being paid out and released participants were asked to fill out

a questionnaire that gathered information on their socio-economic characteristics.

4.3 Predictions and hypotheses

Consider a subject i that is asked to choose how much risk he wants to share with subject j. If subject i is motivated by his own material interest, he will choose a risk sharing level s_i in order to maximize the expected utility of his earnings. Four states of the world k need to be taken into account: both subjects in the pair win, both lose, i wins and j loses, j wins and i loses. Formally stated:

$$\max_{s_i} EU_i = \sum_{k=1}^{4} p_k \cdot U(\Pi_{i,k}) \tag{4.2}$$

where p_k indicates the probability of the state of the world k and $U(\cdot)$ is the utility of i's final earnings, $\Pi_{i,k}$, in state k (see Equation 4.1). The optimal risk sharing level s_i^* depends on the decision maker's risk preferences, as captured by the shape of $U(\cdot)$, and on the risk exposure of both individuals in the risk sharing pair. Hence, if subjects are self-interested, we should observe no significant difference in risk sharing in ENDO and EXO when risk exposure and risk preferences are taken into account. Further, the risk exposure of the sharing partner should influence risk sharing decisions in the same way in EXO and ENDO.

Abundant empirical evidence has demonstrated that, in contrast to the classical assumption of self-interested agents, a considerable fraction of individuals are characterized by a concern for others (see, for example, Camerer, 2003 and Forsythe et al., 1994). Moreover, many individuals are willing to support some degree of redistribution in favor of the less fortunate, even at a personal cost (see Fong, 2011, and Bowles and Polania-Reyes, 2012). In our experiment, the more risk is shared the more eventual income differences between sharing partners are reduced. Hence, individuals' decisions to share risk may be influenced by their distributional preferences, with individuals sharing more the stronger their preference for equality. Among other factors, the strength of distributional preferences has been found to depend on the process that generates income. In particular, inequalities due

to factors within individuals' control, such as effort, are perceived as largely justifiable, while more redistribution is observed when income differences are attributable to elements beyond peoples' influence, such as pure luck (see Alesina and Glaeser, 2004, Cappelen et al., 2007, Durante and Putterman, 2009 and Krawczyk, 2010). In our set-up, participants can influence their income in ENDO by actively choosing between two lotteries that differ in variance. In each situation, by choosing the riskier lottery they can potentially achieve the highest earnings. However, compared to its alternative, the riskier lottery at the same time entails either an increased likelihood of the bad state or a lower outcome in the bad state. In contrast to ENDO, individuals cannot influence their income in EXO as risk exposure is randomly assigned.

Since risk sharing entails redistribution, everything else equal, individuals' willingness to share risk may be higher in EXO as compared to ENDO. The strength of responsibility attributions in ENDO may however be weak, as risk sharing decisions are made *before* lottery outcomes are known, that is when only choices, but not their consequences, are observed. Thus, our first hypothesis is that average risk sharing is lower in ENDO than in EXO, but we do not expect this difference to be considerably large **[Hp1]**.

Cappelen et al. (2013) find that after risks are resolved, most individuals are not in favor of redistributing income from individuals who avoid risk to high risk takers that got a low outcome. At the same time most individuals are willing to eliminate ex post outcome differences resulting from differences in luck among risk-takers. This shows that individuals' preferences for redistributing income are affected by whether others choose to expose themselves to risk or not. Thral and Rademacher (2009) study how much individuals that choose a safe option are willing to transfer to individuals that instead choose a risk, and loose. The authors compare transfers to a situation in which all participants are exposed to risk, and find that subjects are less generous towards those whose bad outcome is a result of their risk-taking actions compared to those who could not influence their outcome. As this evidence suggests that responsibility for risk taking matters for redistribution preferences, we want to investigate whether responsibility attributions for risk exposure affect risk sharing in a similar manner. In particular, we hypothesize that low risk takers share less with high risk

takers as compared to what they share with low risk takers, and that this difference is smaller in the EXO treatment **[Hp2]**. Employing the strategy method allows us to observe risk sharing decisions both for the case where the partner's risk exposure is high and low. Hence, we can test whether in ENDO individuals condition risk sharing decisions on their partner's risk exposure to a different extent as compared to EXO.

4.4 Results

We start our analysis by comparing average risk sharing between treatments and proceed by investigating risk sharing within treatments. In particular, we investigate how an individual's decision to share risk depends on his risk exposure and risk preferences, and on the risk exposure of the sharing partner. Before we present the results on risk sharing, we estimate participants' risk preferences using the elicited certainty equivalents from the third part of the experiment.

4.4.1 Elicited risk preferences

For the elicitation of participants' risk preferences we assume a power utility function for money $U(x) = x^\alpha$ and estimate the parameter value of $\alpha \in]0, \infty[$ at the individual level, by minimizing the sum of squared distances (see Wakker, 2008 and Wakker, 2010). That is:

$$\min_\alpha \sum_n [(p_n H_n^\alpha + (1-p_n) l_n^\alpha)^{\frac{1}{\alpha}} - ce_n]^2$$

where the first term in brackets indicates the theoretically predicted certainty equivalent for lottery n, and ce_n is the elicited certainty equivalent of lottery $n = 1,..8$. To correct for heteroscedasticity lotteries are normalized to uniform length. We find that the median participant is characterized by $\alpha = 0.89$ (s.d. $\alpha = 0.41$, mean $\alpha = 0.93$). A majority of 67% of participants is risk averse. In the remainder we focus on the results regarding subjects that in the second part of the experiment are selected to choose the risk sharing levels.

Since options are randomly assigned in the EXO treatment, we observe no correlation between subjects' estimated risk preferences and the type of

lotteries assigned to them (Spearman's rho=-0.01, p-value=0.96). In other words, in many cases an individual is exposed to a risk that is not in line with his risk preferences. On the other hand, in ENDO the estimated coefficient of risk aversion α is positively correlated with the number of times an individual chooses the riskier option in the risk exposure part (Spearman's rho 0.42, Pearson correlation 0.40, p-value$<$ 0.05). This implies that the more an individual is risk seeking (averse) the more often he selects the riskier (less risky) lottery in the risk exposure part. Table 4.2 indicates, for each situation, the percentage of risk averse and risk seeking individuals that in the risk exposure part choose a lottery in line with their estimated risk preferences. From now on, we refer to those cases as 'consistent', and as 'inconsistent' otherwise.

Table 4.2: Risk preferences and risk exposure, ENDO

Situation	Choose r and $\alpha < 1$	Choose R and $\alpha > 1$
I	67%	50%
II	70%	39%
III	57%	72%
IV	60%	67%

In a majority of cases subjects choose a risk exposure in line with their estimated risk preferences, the occurrence of consistent behavior being especially high among risk averse individuals.

In the following analysis we present our results on risk sharing. For the interpretation of risk sharing decisions, in the EXO treatment we wish to control for whether subjects are exposed to a risk that is likely to be desirable for them, as this changes the incentives for sharing risk with others. To this end, we use individuals' estimated risk preferences. In the ENDO treatment elicited risk preferences are a good predictor of risk exposure choices: thus, we first conduct the analysis of risk sharing behavior taking only individuals' chosen risk exposure into account. Successively, we also distinguish individuals based on whether their chosen risk exposure is in line with their estimated risk preferences.

4.4.2 Risk sharing

When averaging over all situations, we find that risk sharing is equal to 56% in EXO and lower in ENDO, with 50%. In order to test whether this difference is statistically significant **[Hp1]**, we run an OLS regression with risk sharing as the dependent variable, controlling for the risk exposure of both sharing partners and the decision maker's estimated risk preferences. We find that the coefficient of the treatment dummy is insignificant, and thus conclude that there are no significant differences in risk sharing levels between the ENDO and EXO treatment. All regression results are reported in the Appendix.

Result. *Average risk sharing is not significantly different when risk exposure is random as compared to when it is an individual choice.*

In order to test our second hypothesis **[Hp2]**, in the remainder we focus on the relation between risk sharing decisions and risk preferences. Since we employ the strategy method, we can apply a within-subject analysis. We compare individuals' risk sharing behavior when their sharing partner faces option r to the case where their partner faces the riskier option R.

We start by distinguishing individuals based on their risk exposure only and first look at those that were assigned (EXO) or chose (ENDO) the safer lottery r. Figure 4.1 shows average risk sharing in all situations separately for both treatments. The dark bars display average risk sharing levels in cases where both individuals i and j face option r, whereas the light bars represent average risk sharing when the sharing partner, j, faces option R instead.

Averaging over all situations, we find that in the EXO treatment risk sharing is 55% when the partner's option is r and 57% if it is R. A two-tailed Wilcoxon signed-rank test, henceforth WS, shows that these risk sharing levels are not significantly different (p-value=0.60). Results remain insignificant even when we only consider those subjects characterized by risk averse preferences, as measured by α (WS test p-value$\leq$ 0.24). This is important because it shows that results for EXO are not driven by the presence of individuals who are exposed to a risk that they would likely not have chosen by themselves. A WS test conducted for each situation separately confirms

that when risks are exogenous, the partner's risk exposure is not related to risk sharing in a statistically significant way (WS test p-value≥ 0.51).

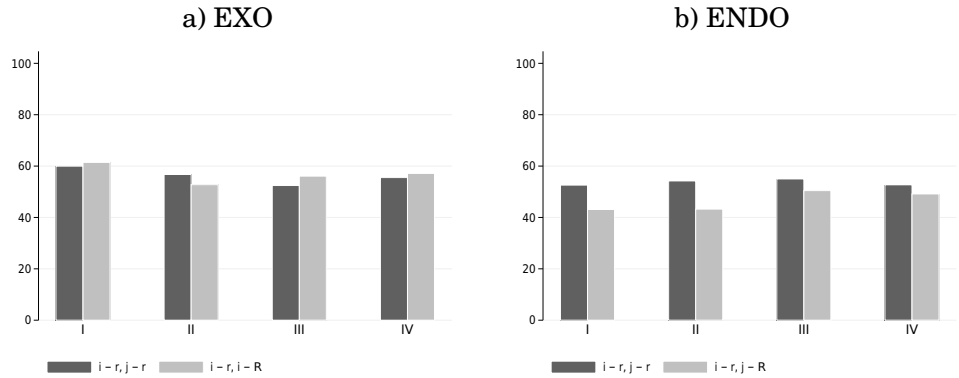


Fig. 4.1: Average Risk sharing of subjects facing option r (%)

Results are fairly different in the ENDO treatment. The average risk sharing level is 54% when the sharing partner chooses lottery r and only 45.5% in case the partner opts for lottery R. This difference is highly significantly different (WS test p-value=0.01). A systematic trend in behavior in all situations can be observed. Among individuals who choose r, less risk is shared on average when the partner chooses option R as compared to option r. The difference is statistically significant in situations I and II, in which high risk taking includes the possibility of a zero outcome (WS test p-values≤ 0.08). Results are not significant in situations III and IV (WS test p-values≥ 0.49). We thus conclude that *only* when risk exposure is deliberate, individuals condition their risk sharing behavior on their partner's risk exposure.

Result. *Individuals who choose to expose themselves to a low risk share less risk with a high risk taker as compared to someone who also makes a cautious choice. This difference is statistically significant when high risk taking includes the possibility of a zero outcome. When risk exposure is randomly assigned, individuals do not condition their risk sharing on their partner's risk exposure.*

We now conduct the same type of analysis for individuals that are assigned (EXO) or choose (ENDO) the riskier option R. We find that in both

treatments and in all decision situations, individuals do not condition their risk sharing choices upon the risk exposure of their partner (WS test p-value≥ 0.25).[5] Average risk sharing in EXO is 53% when the partner is assigned r and 58% when the partner's option is R. The according values are 51% and 53% in the ENDO treatment.

Result. *Individuals that are exposed to a high risk neither condition their risk sharing on their partner's risk exposure when risk is endogenous nor when it is random.*

As anticipated in the previous section, our analysis proceeds by separating the cases where subjects choose (are assigned) a risk exposure in line with their estimated risk preferences from those where this is not the case. This is especially interesting in order to understand risk sharing decisions in the ENDO treatment, as we find that subjects' behavior is systematically related to the consistency of the chosen risk exposure with the estimated risk preferences. This approach confirms that in the EXO treatment risk sharing decisions are never conditioned on the partner's risk exposure. Indeed, we do not observe any systematic trend both at aggregate level (WS test p-value≥ 0.41), as well as at the situation specific level (WS test p-value≥ 0.11) when we control for whether individuals face an option consistent with their estimated preferences. In the following we exclusively focus our attention to the analysis of risk sharing behavior when risks are endogenously chosen.

We first consider the cases where participants choose an option consistently with their estimated risk preferences. The importance of analyzing risk sharing behavior when choices are consistent is evident if considering that such choices are simply most frequent, as shown in Table 4.2. Further, many individuals display the same attitude towards risk across different domains (see Einav et al., 2012 and Dohmen et al., 2011) and hence, the following results are perhaps our most generalizable ones.

Among risk averse individuals that choose the low risk option r, the average risk sharing level is 58% when the sharing partner also chooses option r, but 10 percentage points lower when the partner chooses option R (WS test

[5] Figure 4.3 in the Appendix shows average risk sharing levels in both treatments.

p-value= 0.04).[6] Figure 4.2 a) shows the average risk sharing levels in each decision situation.

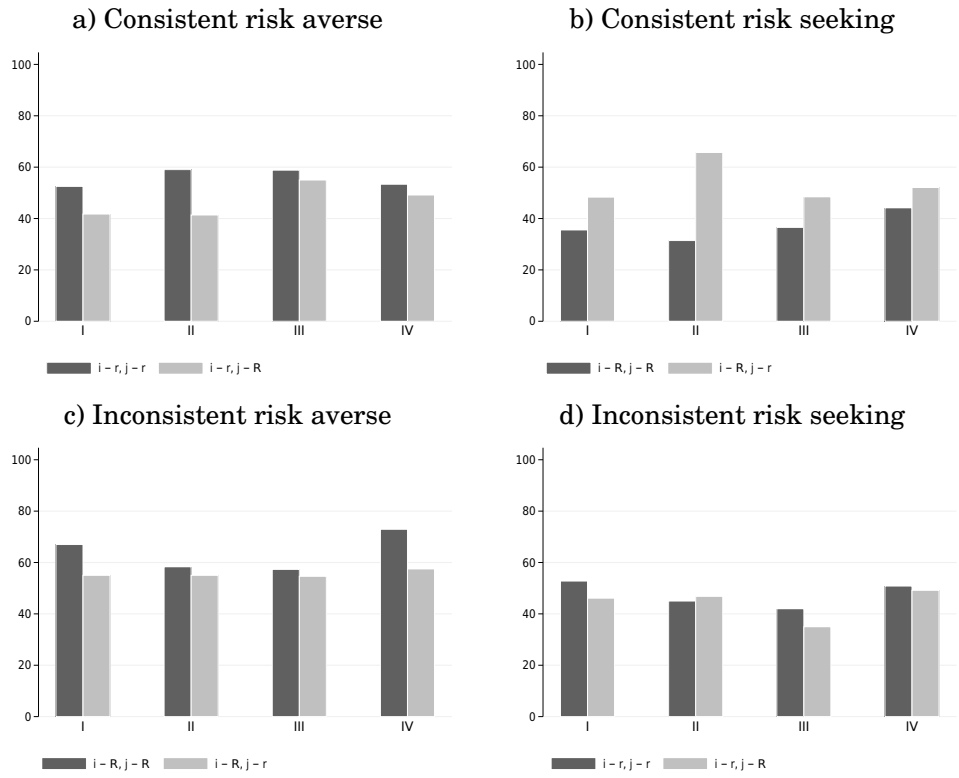


Fig. 4.2: Average risk sharing in ENDO taking consistency into account

In line with the results of the previous section, we find the same pattern of behavior in situations I to IV: i's willingness to share risk with j is lower when j chooses the riskier lottery, R. This difference is statistically significant in situations I and II (WS test p-value= 0.08 and 0.04, respectively). In situations III and IV differences are insignificant (WS test p-value≥ 0.4).

We now consider risk seeking participants that act consistently, and thus choose R. The average sharing level is 42% in case the sharing partner also chooses option R. In contrast, in case the sharing partner chooses option

[6] Our unit of observation is an individual's average risk sharing level calculated over those situations where the individual chose an option in line with his estimated risk preferences.

r, the average sharing level is 51% (WS test p-value= 0.11). Figure 4.2 b) shows that when risk seeking subjects act consistently, they systematically share more risks with subjects who choose the safe option. Differences are statistically significant at the 5% level in situation II. In the other situations, most likely because of the limited number of observations, differences are insignificant (WS test p-value≥ 0.16). In sum, risk sharing behavior of consistent risk seeking participants displays the same tendency as in cases where risk averse individuals choose consistently: facing a high risk taker as sharing partner reduces people's willingness to share risks.

Result. *When risk exposure is endogenous and individuals choose their risk exposure in a consistent way, they tend to share more risk with a partner that made a cautious choice as compared to a risky one.*

In the last part of the analysis we turn our attention to the cases where individuals in the ENDO treatment choose an option that is not in line with their estimated risk preferences. We will not try to provide an explanation of the observed inconsistencies, but rather focus on understanding risk sharing behavior in these cases. We begin by considering risk averse individuals who select option R. Risk sharing is on average 61% when their sharing partner also chooses option R. On the other hand, when the other member chooses option r, the average sharing level is 50%. Hence, in contrast to what is observed when risk averse individuals act consistently, they share less risk if their partner chooses option r (WS test p-value=0.14). Figure 4.2 c) shows the average sharing levels per situation. Note that in all situations i shares on average less risk when j chooses the safer option r as compared to when he chooses R. This difference is statistically significant in situation I and IV (WS test p-value=0.08 and 0.09 respectively). In the other situations differences are statistically insignificant (WS test p-value≥ 0.39). We conclude that in case generally cautious individuals select the riskier option they prefer to share more risk with individuals who make the same risky choice.

Result. *Risk averse individuals that decide to expose themselves to a high risk tend to share less risk with a cautious sharing partner as compared to a high risk taker.*

Lastly, we note that the behavior of risk seeking individuals that choose inconsistently does not present any systematic trend. The average sharing level is 46% when the sharing partner also chose option r and 45% when he chose option R (WS test p-value=0.78). There are no significant differences within situations (WS test p-value≥ 0.20), as suggested by Figure 4.2 d).

To summarize, we find a lower willingness to share risk with high risk takers whenever individuals choose a risk exposure in line with their estimated risk preferences. However, when individuals deviate from their general risk preferences, they either do not condition their risk sharing on the partner's risk exposure or share more with high risk takers. Importantly, we find that none of the described effects emerges in the EXO treatment, that is when risk exposure is beyond individuals' control. Our results thus allow to conclude that responsibility for risk exposure matters for individuals' risk sharing decisions.

4.5 Discussion and conclusion

In this study we experimentally investigate how individuals' support for risk sharing is related to whether risk exposure is deliberate or unswayable, to individuals' own risk preferences and to the risk exposure of the sharing partner.

Our main result is that when risk exposure is deliberate individuals are less willing to share risk with high risk takers as compared to low risk takers. When risk exposure is instead exogenous, risk sharing decisions are not conditioned on the risk exposure of the sharing partner. The observation that low risk takers' sharing decisions are affected by responsibility attributions carries important practical consequences for voluntary risk sharing arrangements. Indeed, in expectation, individuals who, for example, avoid unhealthy habits contribute more than high risk takers to the risk sharing pool. Hence, their support of risk sharing arrangements is fundamental for their sustainability. Interestingly, even in situations where individuals consistently take high risk they prefer to share more risk with those who choose a low risk exposure.

When averaging over all the different situations we analyze, we also find that risk sharing is higher, but not significantly so, when risk exposure is

random as opposed to deliberate. This suggests that, compared to ex-post redistribution decisions, ex-ante decisions that carry redistribution effects may be less influenced by whether individuals' outcomes are generated by deliberate risky choices. Another possible explanation is related to the fact that in our set-up responsibility is associated to choice under risk, while previous studies have compared income differences due to luck with those due to effort (see, for instance, Cappelen et al., 2007).

Importantly, our results can only be a lower bound of how much the support for risk sharing is a function of others' risk taking behavior. The riskier options in our set-up are characterized by potential outcomes that are more extreme in a negative and positive way, as compared to their safer alternatives. Thus, sharing risks with a high risk taker can be attractive, since it potentially allows to profit from a high income. However, habits like smoking, overeating and reckless driving hardly have positive externalities for society, as taking a high risk can at best increase the utility of the risk taker.

Taken together, our results suggest that measures that at least partly account for risk takers' responsibility for higher expected benefits from the system, such as raising smokers' health insurance premiums, may be desired. However, this is only the case if risk exposure is perceived as a choice. When risks are exogenous, for any risk preference, average risk sharing levels do not systematically vary with the risk exposure of the risk sharing partner. Thus, it seems that when risk is not a choice, redistribution motives may partly override self-interested insurance purposes. It remains an open question under which circumstances some preferences and behaviors are perceived as more exogenous, and thus less deserving to be condemned, than others (Bossert and Fleurbaey, 1996). An interesting avenue for future research would be to investigate how individuals' perception of choice responsibility can be influenced in order to promote support for risk sharing.

We conclude with a note on those situations where individuals who are generally cautious choose a high risk. The fact that these individuals share more with partners that made the same risky choice suggests that when they 'dare' to take risks they can better identify with other high risk takers. This identification effect may keep individuals from lowering risk sharing and, in contrast, induce them to even share more risk with the partner they identify with. This interpretation is suggested by studies showing a positive

relationship between social identification and willingness to redistribute income (see Klor and Shayo, 2010, and Fowler and Kam, 2007). Further investigations are however needed to test the robustness of this effect.

4.6 Appendix

4.6.1 Regression results and additional figure

Table 4.3: Risk sharing level, treatment comparison.

Variable	**Coefficient**	**(Std. Err.)**
ENDO	-4.753	(3.752)
i-r, j-R	-3.638	(2.331)
i-r, j-R	-1.139	(3.396)
i-R, j-r	-2.018	(2.843)
alpha	1.758	(3.910)
Intercept	55.860***	(4.402)
N	832	
R^2	0.009	
$F_{(5,103)}$	.969	

Note: Standard errors are robust to heteroskedasticity and are clustered around subjects; the case i-r, j-r is the omitted treatment category; ENDO is a treatment dummy.

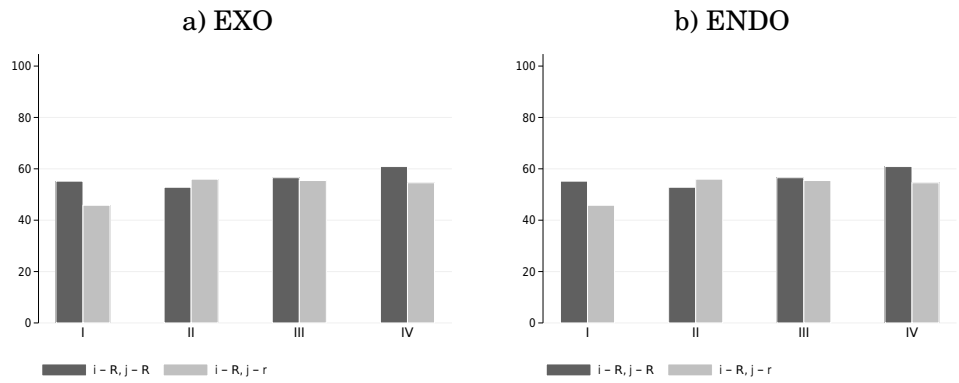


Fig. 4.3: Average Risk sharing of subjects facing option R (%)

4.6.2 Experimental instructions

The following instructions appeared on participants' computer screens (exempt from the headlines in squared brackets).

Part 1 *[Risk exposure part]*
This is a screen shot of a typical decision situation that you are going to face. You are not asked to make choices now! Please have a careful look.

Fig. 4.4: Screenshot risk exposure part

In this example, Option A yields 30 Euro with 20% chance and 0 Euro with 80% chance. Option B yields 12 Euro with 50% chance and 0 Euro with 50% chance. In order to choose between Option A and Option B you will have to tick one of the boxes surrounded by the red frame.
Assume, for instance, that the decision situation above is selected to be relevant for your payment and also assume that you chose Option B. It follows that at the end of the experiment a random draw will determine whether your outcome is 12 Euro or 0 Euro.
This outcome, together with the other decisions that you will make in this part of the experiment will determine your final earnings. In the following screens we will explain in detail how your earnings are affected by this outcome.

Part 1 (continued) *[Risk sharing part]*
After everyone has made choices between Option A and Option B, you will be randomly matched with another participant in the room. The two of you form a group. One person in the group (you or the other participant) will be randomly selected. For each of the 5 decision situations you faced before, the selected person has to choose a number between 1 and 100, which determines the individual outcome's percentage that each group member deposits in a group account.
At the end of the experiment, one decision situation will be randomly selected and the outcomes of the chosen options will be determined. Consequently, the amount in the group account will be calculated and equally divided between the two persons in the group. Notice that the selected person is asked to choose the percentage before the outcomes of the chosen options are known.

Assume, for instance, that you are selected to choose the percentages. In a given decision situation, you will have to make a choice before you know the outcome of the option you chose and before you know the outcome of the option chosen by the other person in the group.
Also notice that the selected person has to choose percentages for each of the 5 decision situations because the decision situation relevant for payment is only determined at the end of the experiment. Therefore, each choice has

to be considered in isolation from the others, as if it were the one which is relevant for payment.
On the following screen the determination of earnings is illustrated with the help of examples.

Part 1 (continued)
Imagine now that you have been selected to choose the percentage.

Example 1. Assume that both you and the other person in the group have chosen Option A, which in the preceding example yields 30 Euro with 20% chance and 0 Euro with 80% chance. Let's say that you choose the percentage value 40%. At the end of the experiment the uncertainty concerning your earnings is resolved; assume that Option A eventually yields 30 Euro to you and 0 Euro to the other person. It follows that:
- The group account consists of 12 Euro (that is, 0.4*30+0.4*0=12+0).
- Your return from the group account is 6 (=12/2) Euro.
- Your earnings are 24 Euro (=30-12+6).
- The earnings of the other person in the group are 6 Euro (=0-0+6).

Example 2. Assume now that in the same decision situation you have chosen Option A, which yields 30 Euro with 20% chance and 0 Euro with 80% chance, and that the other person chose Option B, which yields 12 Euro with 50% chance and 0 Euro with 50% chance. Let's say that you choose the percentage value 70%. At the end of the experiment the uncertainty concerning your earnings is resolved. If Option A eventually yields 30 Euro to you and Option B yields 12 Euro to the other person. It follows that:
- The group account consists of 29.4 Euro (that is, 0.7*30+0.7*12=21+8.4).
- Your return from the group account is 14.7 (=29.4/2) Euro.
- Your earnings are 23.7 Euro (=30-21+14.7).
- The earnings of the other person in the group are 18.3 Euro (=12-8.4+14.7).

This is a screen shot of a typical decision situation that you are going to face. You are not asked to make choices now! Please have a careful look.
In case you are selected to choose the percentages, you are asked to do this twice for each decision situation. First, for the case that the other person in your group chose Option A (red circle) and second for the case that the other person in your group chose Option B (red square). When choosing the percentages you will not be informed about the actual option chosen by

Fig. 4.5: Screenshot risk sharing part

the other person in your group. At the end of the experiment the percentage associated to the actual choice of the other person in your group will be implemented. In other words, you will choose percentages for two possible scenarios. Since you do not know which one will be relevant for your payment, you have to make each choice in isolation and with the same accuracy. Notice that your outcome may be different from that of the other person in your group even if both chose the same option.
Also notice that if you choose a percentage of 0 your earnings and the earnings of the other person in the group will exclusively depend on the individual outcome of the option that each of you chose. Conversely, if you choose a percentage of 100 your earnings and the earnings of the other person in the group will be equal to each other, as they will be the sum of your outcomes divided by 2.

Part 2 *[Elicitation of risk preferences]*
You are now going to make a series of decisions. These decisions will not influence your earnings from the first part of the experiment, nor will the decisions you made in the first parts of the experiment influence the earn-

ings from this part. Furthermore, the decisions you are going to make will only influence your own earnings.
You will be confronted with 9 decision situations. All these decision situations are completely independent of each other. Each decision situation is displayed on a screen. The screen consists of 20 rows. You have to decide for every row whether you prefer Option A or Option B. Option A is the same for every row in a given decision situation, while Option B takes 20 different values, one for each row.
Note that within a decision screen you can only switch once from Option B to Option A: if you switch more than once a warning message will appear on the screen and you will be asked to change your decisions.
This is a screen shot of a typical decision situation that you are going to face. You are not asked to make choices now! Please have a careful look.

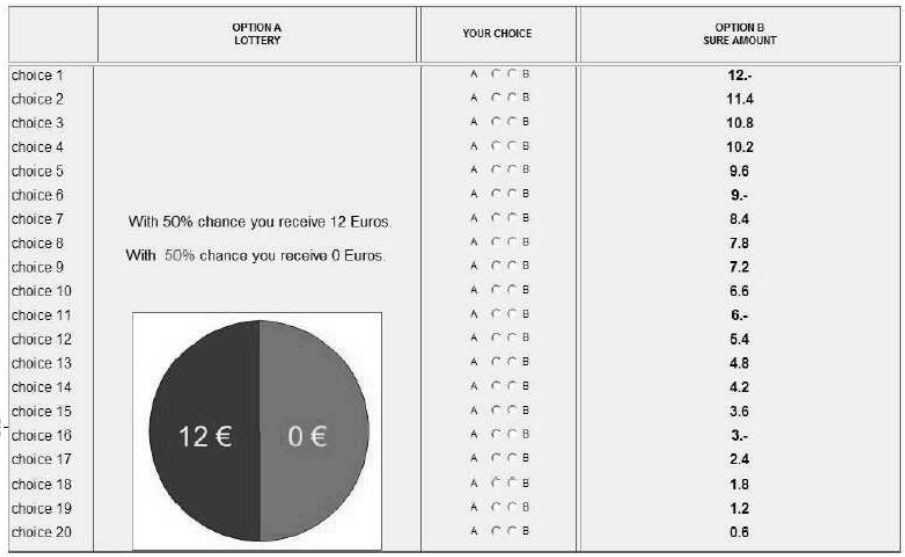


Fig. 4.6: Screenshot risk preferences elicitation

Determination of earnings
At the end of the experiment one of the 9 decision situations will be randomly selected with equal probability. Once the decision situation is selected, one

of the 20 rows in this decision situation will be randomly selected. The choice you have made in this specific row will determine your earnings.
Consider, for instance, the first screen shot that you have seen. Option A gives you a 50% chance to earn 12.- Euro and a 50% chance to earn nothing. Option B is always a sure amount that ranges from 12.- Euro in the first row, to 0.6 Euro in the 20th row. Suppose that the 12th row is randomly selected. If you would have selected Option B, you would receive 5.4 Euro. If, instead, you would have selected option A, the outcome of the lottery determines your earnings. At the end of the experiment the lottery outcome will be determined by the computer.

Please note that each decision situation has the same likelihood to be the one that is relevant for your earnings. Therefore, you should view each decision independently and consider all your choices carefully.

Chapter 5

The Benefits of External Control

5.1 Introduction

In principal-agent relationships the benefits of the principal typically depend on the extent to which the agent exerts costly effort. It is thus in the principal's interest to control the agent in order to ensure a minimum benefit for himself. Whereas standard economic theory evaluates controlling interventions to be effective in increasing effort, there is evidence that control can have detrimental effects on agents' effort provision. Falk and Kosfeld (2006), henceforth referred to as FK, implement a principal agent game where the principal can decide to control the agent by imposing a minimum effort requirement before the agent makes his effort decision. The authors find evidence for hidden costs of control in the sense that a majority of agents transfer less to the principal if controlled than if not controlled. In their experiment these hidden costs of control even outweigh the disciplining effects that control has for the selfish agents with the consequence that effort provision is lower on average when control is imposed as compared to when the principal refrains from controlling.

A multitude of mechanisms have been proposed to rationalize this negative effect. The constraint may be interpreted as a signal that the principal is selfish and is thus not worth impressing with high effort provision (Ellingsen and Johannesson, 2008), that the principal is pessimistic about the social norm of effort provision and believes that most agents are selfish, which in turn leads conformist types to behave selfishly (Sliwka, 2007) or that the controller has low expectations about the agent's effort provision, which some agents perceive as distrusting (FK). Also, "not being controlled might be considered to be kind, because not everybody reciprocates not being controlled with high effort", which leads the agents to positively reciprocate with high effort (Siemens, 2013, p.55). Being controlled may further be disliked as it constrains the agent's choice autonomy (FK) or control may undermine agents' intrinsic motivation to behave pro-socially (see e.g. Deci, 1971; Gneezy and Rustichini, 2000).[1]

Although many studies find evidence for hidden costs of control in the set up of FK, most of them do not find that the costs outweigh the disciplining effect that control has on the selfish agents (see Schnedler and Vadovic,

[1] The different mechanisms are discussed in more detail in Section 5.2 and Section 5.5.

2011, Kessler and Leider, 2013, Hagemann, 2007, Ziegelmeyer et al., 2012). The overall effect of control on agents' effort provision might be negative, however, if negative spillover effects on the amount of effort provided to other principals exist. Constraining an agent's choice set may not only lower his intrinsic motivation to behave pro-socially towards the principal whose transfer is affected by the constraint, but it may lower the agent's intrinsic motivation to behave pro-socially towards *any* principal. Spillover effects may also occur if control is indeed interpreted as a signal that low effort provision towards principals is the social norm among agents (Sliwka, 2007). In order to test whether such indirect hidden costs of control exist, we extend the experimental setting in FK to a setting with two principals and one agent. Individuals are randomly matched in groups of three and assigned the role of an agent A or one of two principals B and C. The agent may transfer up to 100 points to each of the two principals who do not have any endowment. Transfers x_B and x_C are costly to him but beneficial for the respective principal. Before the agent makes his decision, principal B may decide whether to impose a minimum transfer $\underline{x}_B > 0$. The size of $\underline{x}_B$ is varied between treatments and is either 5 or 10. Importantly, the choice set for x_C is not affected by the constraint.

We investigate two different types of control. In the internal control treatment the earnings of the controlling principal, henceforth called 'controller', depend on the transfer that is constrained. Control is thus internal to one of the principal agent relationships. In our external control treatment another person than the principal himself decides about whether to impose $\underline{x}_B$. The controller's earnings are thus independent of the agent's decision. Hidden costs of control in the latter set-up cannot be explained by a mechanism that works via a negative reaction towards the controller, such as negative reciprocity, punishment of distrust, or a lower desire to appear pro-social to the controller. Hidden costs of imposing a constraint that is perceived as controlling may nevertheless occur via a crowding out of intrinsic motivation to behave pro-socially (Deci, 1971, Gneezy and Rustichini, 2000, Gneezy et al., 2011). Also, trusting behavior by an external person may, likewise the case of internal control, be interpreted as a signal that the social norm is to behave pro-socially which is then imitated by conformist agents (Sliwka, 2007).

In summary, the aim of this paper is twofold: first, we investigate whether control (internal or external) entails hidden costs of control via the constrained transfer and second, we investigate whether spillover effects to a second transfer exist in which control can not be applied. Eliciting the agents' transfer decisions for both possible decisions of the controller (control versus not control) allows us to determine the fraction of agents reacting negatively, neutrally or positively to control towards the two principals. If the existence of hidden costs of control hinges on the direct impact that the constrained agent has on the controller's earnings, external control can be an effective way to reap the benefits of control without provoking detrimental side effects.

We find that with internal control, a considerable fraction of agents provide less effort towards their principal if he decides to control as compared to if no control is imposed. Their unconstrained effort provided to the second principal is, however, not systematically affected by the control decision. The overall amount of effort provided by the agents to both principals does not depend on the controller's decision. In contrast, effort provision is significantly higher for both principals when control is imposed by an external controller than when there is no control. For a low minimum transfer requirement of 5 we find that the transfers to principal B when control is imposed are significantly higher when this is done by an external controller as compared to an internal one.

Our results suggest that control per se does not necessarily provoke negative reactions. Instead, the occurrence of detrimental effects depends on who is the controller, i.e. whether the controller is the principal who profits (suffers) from high (low) effort provision by the agent or an external entity whose earnings are not at stake. Our finding that external but not internal control is effective in increasing agents' effort offers a rationale for the delegation of the control decision. Existing literature shows that in the context of distribution games, responsibility for decision outcomes can be turned away from a first mover who shifts the distribution decision to a third party (see e.g. Bartling and Fischbacher, 2012, Coffman, 2011). This is the case even if the first mover can influence the decision that is eventually made. Likewise, in our context, less individuals may potentially react negatively to control if principal B delegates the control decision to an external party.

5.2 Related literature

As shortly reviewed before, several theories attempt to explain the phenomenon observed in FK. Ellingsen and Johannesson (2008) provide a theoretical social-esteem model, in which pro-social agents have a stronger desire to make a good impression on a pro-social principal as compared to a selfish one. Thus, depending on the agent's expectation about the principal's type his incentive to appear pro-social varies. When the principal decides to impose a constraint on the agent's transfer he signals that he is a selfish type whereas refraining from control signals that he is a social type. Siemens (2013) explains the phenomenon of hidden costs of control by intention based reciprocity. The author assumes that individuals differ in their propensity for reciprocity and that worker's preferences are private information, i.e. the principals do not know the shares of reciprocal and selfish workers. If a principal decides not to control, this is kind towards the selfish individuals as it allows them to reduce their effort and thus increase their payoffs. At the same time it is unkind towards the reciprocal agents as they react by increasing effort which however implies a lower payoff for them. If agents however take into account that principals legitimately believe them to be selfish, not imposing control can also be perceived as kind action by reciprocal agents, leading them to reciprocate with higher effort. Sliwka (2007) provides a signaling model, in which the principal's control decision signals his confidence that many workers are selfish. The author assumes that for one type of workers, the conformist workers, preferences are determined by their belief about the relative frequency of selfish and fair-minded workers. They adapt the preferences of those individuals that they believe are in the majority and thus behave more selfishly when controlled. The model is based on the assumptions that some uncertainty about appropriate behavior exists and that the principal's decision gives a credible signal about the distribution of types.

Our experiment is further related to studies on the effects of monetary incentives via an interaction with psychological incentives (see Gneezy et al., 2011, Frey and Jegen, 2001 or Bowles and Polania-Reyes, 2012 for overviews). For example, Gneezy and Rustichini (2000) find that low monetary incentives have a detrimental effect on performance for Israeli high

school children who are doing volunteer work (collecting monetary donations). Introducing a performance dependent payment leads to significantly lower amounts collected as compared to when no monetary compensation was offered. The findings suggest that the childrens' reward was shifted from social approval to receiving money which was a lower incentive for them to perform well. Boly (2011) implements experiments in the field and in the laboratory using a real-effort task which consists of finding mistakes in papers. The number of undetected mistakes determines the size of the penalty that is imposed on the participants. Comparing two monitoring treatments that differ in their severity to a benchmark of no monitoring, he finds that monitoring significantly increases effort. This difference is independent of the severity of monitoring. He further finds that monitoring only has a significant disciplining effect on selfish graders, but not on intrinsically motivated ones. The grader's decision to reject a bribe offer by one of the authors of the graded papers serves as a proxy for intrinsic motivation. Fehr and Rockenbach (2003) conduct a trust game in which the investor has the possibility to indicate a desired back transfer and to impose a fine on the trustee in case he does not meet his claim. The authors compare behavior to a trust game in which the imposition of a fine is not possible. They find that if the investors impose the fine, trustees' back- transfers are smaller as compared to when no fine option existed or when the fine was not imposed. Moreover, they find that back-transfers are significantly lower when the desired back-transfers are high than when they are low. They conclude that economic incentives that are associated with greedy or selfish intentions cause negative effects on altruistic cooperation. As Gneezy et al. (2011) point out a "basic condition for the existence (and empirical evidence) of intrinsic motivation is that the activity should be exercised even when reward is absent". Similarly, in our experiment, some agents may be intrinsically motivated to behave pro-socially towards the principal even when no control is imposed. In that case control may have an adverse effect on individuals' motivation to behave in the principals' interest.

FK argue that the feeling of distrust and a perceived lack of choice autonomy are responsible for the existence of hidden costs of control. They conduct a control treatment in which they eliminate distrust and control by presenting the participants with a game in which the principal does not decide whether

to control or not (nor does another person), but instead the strategy set of the agent is restricted right away. As no active choice decision of the experimenter is made (salient), participants may not feel controlled or distrusted. The authors do not find a crowding out effect in that setting.[2]

Concerning the existence of spillover effects, to the best of our knowledge, only principal-agent relationship *internal* spillover effects have been analyzed so far. This means spillover effects from a constrained effort dimension to another unconstrained one while both dimensions affect the same principal's earnings. Frey and Benz (2001) investigate experimentally what they call the 'motivation transfer' effect. This effect implies that incentives reduce intrinsic motivation in areas that are not subject to the intervention. They conduct a trust game experiment in which an employer first transfers a wage to the agent, indicates a desired quantity (back transfer) and can then decide whether to impose a conditional deduction on the agent in case he does not meet his desires. Then the agent decides how much work quantity *and* how much work quality to provide. The authors find that the imposition of the conditional deduction also crowds out intrinsically motivated voluntary cooperation in the area not subject to the restriction, the work quality. Belot and Schröder (2013) conduct a field experiment in which they investigate the effect that monitoring has on various forms of counterproductive behavior. They find that monitoring in one dimension has negative spillover effects to other unmonitored dimensions where cheating is least costly. In our experiment we investigate spillover effects *between* principal agent relationship, i.e. we investigate whether the imposition of control not only affects the transfers to the principal that is affected by the constraint, but whether it also affects the agents' uncontrollable transfers to another principal.

Several experimental papers replicate the existence of hidden costs of control in the FK set-up and investigate the conditions under which hidden costs indeed occur. Schnedler and Vadovic (2011) make control more legitimate in their experiment by a) introducing the possibility that the principal

[2]The authors further find that behavior does not differ depending on whether individuals make their decisions after they learn about the principal's choice or whether the strategy method is applied. Also, they find the existence of hidden costs of control to be robust to a gift-exchange setting where the principal can additionally set a wage level beforehand.

faces a selfish computerized agent who determines the effort of the agent and b) changing the frame of the agent's and the principal's decisions such that the possible restriction prevents stealing from the principal. They find that a crowding out effect does not exist when control is legitimized because it prevents selfishness or theft. Kessler and Leider (2013) find that hidden costs only exist if a) a high effort-provision norm preexists, b) control is imposed unilaterally and has an asymmetric effect on the agent, c) control is weak and d) the agent would also control if he was in the role of the principal. Ziegelmeyer et al. (2012) observe hidden costs of control that outweigh the disciplining benefits of control when decisions are made hypothetically and are thus not relevant for participants' earnings. When participants across several subject pools are paid based on actual decisions made in the experiment a crowding out effect can be observed, but it is not found to dominate the benefits of control. Lastly, Hagemann (2007) find that the existence of hidden costs of control is highly sensitive to the framing of the experimental instructions. Describing the imposition of the minimum transfer as 'the principal is able to *force* the agent' leads to significantly more hidden costs of control as compared to 'the principal is able to *constrain* the agent' (the latter wording we apply in our instructions).

The remainder of the paper is organized as follows. Sections 5.3 and 5.4 describe the experimental design and procedures. Section 5.5 summarizes theoretical predictions and hypotheses. Results are presented in Section 5.6. In Section 5.7 we discuss the results and conclude.

5.3 Design

We implement the following multiple principals-agent game in the experimental laboratory. Individuals are randomly matched in groups of three and assigned the roles of either the agent A or one of two principals B and C. The agent is endowed with 220 points, whereas the principals B and C do not have any endowment. The agent chooses effort levels x_A and x_B, which are costly to him but beneficial for the respective principal. The costs for the agent are $c(x_B) = x_B$ and $c(x_C) = x_C$ while the benefits are $2 \cdot x_B$ for principal B and $2 \cdot x_C$ for principal C. Consequently, the earnings functions are given

by:

$$\Pi_A = 220 - x_B - x_C \text{ for agent A}$$
$$\Pi_B = 2 \cdot x_B \text{ for principal B}$$
$$\Pi_C = 2 \cdot x_C \text{ for principal C}$$

Before the agent decides on x_B and x_C the agents' choice set for the transfer x_B may be constrained. The person eligible to impose this constraint is called the 'controller'. He can either impose a fixed minimum transfer $\underline{x}_B > 0$ for the agent's transfer to principal B such that $x_B \in [\underline{x}_B, ..., 99, 100]$ or he can leave the choice set unconstrained, allowing the agent to choose any integer value between 0 and 100. The choice set for x_C can not be constrained such that $x_C \in [0, 1, 2, ...100]$ always holds. We chose this one-dimensional constraint set-up in order to investigate whether the imposition of a constraint $\underline{x}_B$ has an impact on the unconstrained transfer to principal C.

We implemented four treatments that differ with respect to *who* is given the option to constrain the agent and with respect to the *size* of $\underline{x}_B$. In the two internal control (IN) treatments, principal B is the controller. Thus, the agent's eventual reaction on being constrained or not directly affects the earnings of the controller. In the two external control (EX) treatments, the controller is a person outside the group. Specifically, this will be a principal B* of another group. B*'s earnings do not depend on A's transfer decision but on the transfer decision of agent A* in his own group (who may be constrained by a principal B' from yet another group). Figure 5.1 displays the interactions structure of the IN and the EX treatments.

In contrast to the IN treatments, in the EX treatments the agent's transfer to principal B that is affected by the constraint does not determine the earnings of the controller.[3] Comparing agents' behavioral reactions to the two types of control allows us to investigate whether the agents' possibility to directly react towards the controller is crucial for the existence of hidden costs of control and the potential spillover effects.
We implement two variations of the IN treatment and the EX treatment respectively, with $\underline{x}_B$ being either 5 or 10 points. The four different treatments

[3]Each participant B is asked to indicate whether he wants to constrain participant A' of another group in his transfer to the principal B' in that other group. See Appendix 5.8.3 for the instructions.

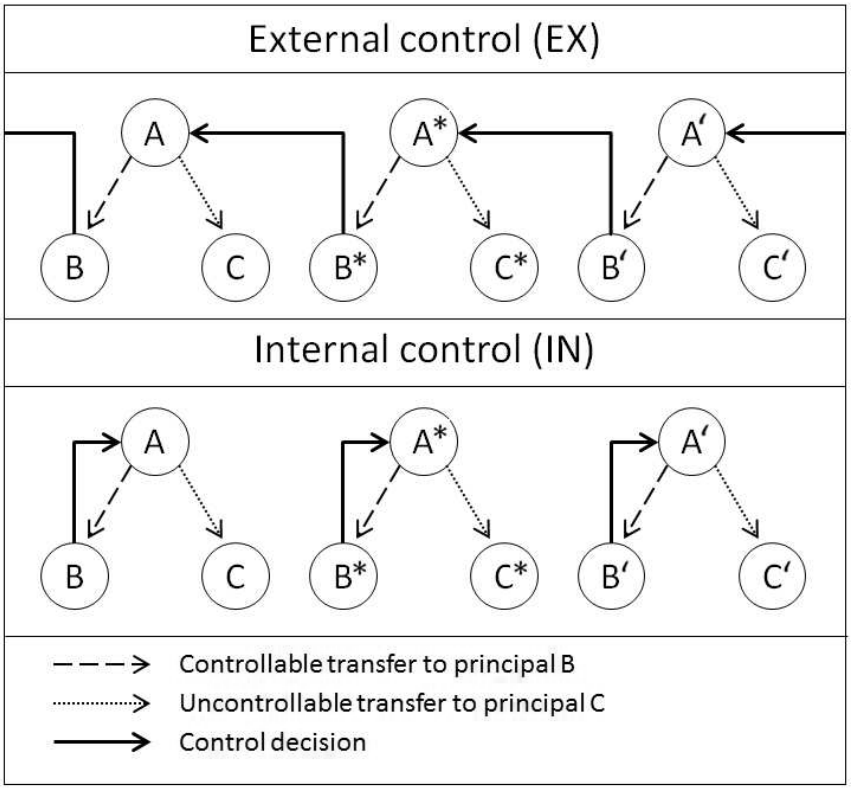


Fig. 5.1: Interaction structure IN and EX treatments

are henceforth denoted as treatments IN5, EX5, IN10 and EX10. Predictions for the agents' behavior will be discussed in Section 5.5.

5.4 Procedures

In all treatments, principals and agents interact only once in the computerized experiment. We employ the strategy method in order to elicit the agents' transfer decisions for two conditions (Selten, 1967). The first one is that the controller imposes the minimum transfer $\underline{x}_B$ (constrained condition) and the second is that he refrains from doing so (unconstrained condition). To control for possible order effects we randomize the order of appearance of the

two cases on the computer screens.[4] For simplification, in the following we will refer to the transfers to principal B and C in the constrained condition as x_B^c and x_C^c with the superscript 'c' standing for 'constrained'.[5] Likewise transfers made in the unconstrained condition are referred to as x_B^{nc} and x_C^{nc} with the superscript 'nc' standing for 'not constrained'. The strategy method allows us to elicit x_B^c, x_C^c and x_B^{nc}, x_C^{nc} for each agent. Which of the two sets of transfer decisions is eventually relevant for participants' earnings depends on the actual decision of the controller. Participants are told that at the end of the experiment a final screen will inform them about the decisions that were made (the controller's and the agent's decisions) and about the resulting earnings in the own group.

In order to increase participants' understanding of the instructions and the earnings procedure they were asked to create and correctly answer three hypothetical control questions administered before the actual start of the experiment.

Before being paid out and released participants were asked to fill out a set of questionnaires (see Appendix 5.8.4). The first questionnaire is related to how individuals perceive the imposition of a constraint and the controller.[6] The second questionnaire is adopted from Dohmen et al. (2009) and elicits individuals' positive and negative reciprocity attitudes. The third questionnaire captures individuals' perception of justice and is adopted from Dalbert (2000). Lastly, individuals were asked to fill out a questionnaire that gathered information on their socioeconomic characteristics.

The experiment was conducted at the Behavioral and Experimental laboratory (BEElab) at Maastricht University. Subjects were recruited on line with the system ORSEE (Greiner, 2004). For the computerized implementation we used the software z-tree (Fischbacher, 2007). Upon arrival at the lab subjects were randomly matched in groups and allocated the role of a principal or that of the agent. A session lasted approximately 35 minutes and the average earnings were 6,50 Euro. In total 300 subjects participated in the

[4] See Appendix 5.8.2 for the decision screens of the controller and the agent.

[5] Note, that 'constrained' is solely referring to the constraint on the transfer to principal B.

[6] Participant A is asked the questions stated in the Appendix, whereas the two principals are asked about their beliefs about A's answers.

experiment (75 in each treatment), out of which 100 participants were assigned the role of an agent. Importantly, each participant only participated in one of the four treatments.

5.5 Predictions

In our experimental set-up there are two decisions that may be affected by the principal's control decision. We will refer to the effect on the controllable transfer to principal B as 'direct' effect, whereas the effect on the uncontrollable transfer to principal C is denoted as 'indirect'.

Selfish agents Irrespective of whether an internal principal B or an external principal B* has the option to impose a constraint on the transfer to principal B, the predictions are identical for purely self-interested individuals. If no constraint is imposed, agents are predicted to not transfer anything to any of the two principals $x_B^{nc} = x_C^{nc} = 0$. If the constraint is imposed they are predicted to transfer the minimum possible to both principals, which is $x_B^c = \underline{x}_B$ and $x_C^c = 0$. Anticipating this behavior participant B always imposes the constraint in the IN treatment. In the EX treatment B* is indifferent between imposing and not imposing the constraint, since this decision does not have any monetary consequences for himself.

Social preferences Abundant empirical evidence has however demonstrated that, in contrast to the classical assumption of self-interested agents, a considerable fraction of individuals can be characterized by a concern for others (see, for example, Camerer, 2003, and Forsythe et al., 1994). Emanating from the existence of outcome-based social preferences, agents are predicted to make positive transfers to principal B and to principal C by choice. That way, they ensure positive earnings for the other participants and reduce inequalities between themselves and others.
If an agent's preferred transfer to B is larger than the minimum required transfer $\underline{x}_B$, outcome-based models predict his choice not to be affected by whether a constraint is imposed or not. Likewise, the transfer to C will always be the same irrespective of the controller's decision. If, instead, the agent's preferred transfer is smaller than $\underline{x}_B$, it is automatically increased to the minimum in case a constraint is imposed and stays unaffected oth-

erwise. If agents have a preference for earning a certain relative share of the payoffs, the transfer to C will be lower in the constrained condition as compared to the unconstrained one (see Bolton and Ockenfels, 2000).

Behavioral reactions to control Agents that are intrinsically motivated to pass on some of their endowment might however show a negative *direct* reaction to the *internal* imposition of a constraint. Potential reasons for this behavioral reaction are discussed in Section 5.2. Control might signal that the controller is selfish and agents have no incentive to appear pro-social to a selfish controller (Ellingsen and Johannesson, 2008), the controller's control decision may be perceived as unkind and lead agents to negatively reciprocate to him by lowering effort provision (Siemens, 2013), or control might signal that the controller has low expectations about the agent's effort provision, which some agents perceive as distrusting and 'punish' with low effort (FK). For all these mechanisms to apply it is necessary, that the controller benefits (suffers) from higher (lower) effort provision by the agent, as it is the case in the internal control treatments. Alternatively, control may give a signal about the principal's confidence that many agents are selfish which leads conformist types to behave selfishly themselves (Sliwka, 2007) or it may interact with agents' intrinsic motivation to behave pro-socially and eventually result in less pro-social behavior (see e.g. Gneezy and Rustichini, 2000). These two mechanisms may also be at play when control is external.

We hypothesize that in line with previous research, e.g. FK, we will find a *direct* detrimental effect of control for a considerable fraction of agents in our IN treatments such that $x_B^{nc} > x_B^c \geq \underline{x}_B$. If the consequences of control are sufficiently detrimental, not imposing it may then be more beneficial for the controller, principal B.

When control is imposed by an *external* controller, it is not possible for the agents to influence the controller's earnings by lowering effort. Thus, the above mentioned first set of mechanisms do not predict detrimental effects of control to occur in this set-up. A crowding out of the agents' intrinsic motivation when feeling controlled, however, as well as the imitation of the behavior that is assumed to be exhibited by the majority of agents as theorized in Sliwka (2007) may predict direct hidden costs of control also when control is imposed externally. Since the agent has no possibility to affect the

controller's earnings in that case, a channel to negatively reciprocate or to punish the controller does not exist. If this, however, is crucial for a negative reaction to control by the agents, we may expect that detrimental effects of control are less pronounced (or non-existent) in the EX treatments as compared to the IN treatments.

We now turn to the *indirect* effects of control via the transfer to principal C. We first consider agents that are intrinsically motivated to give more than the minimum constraint to principal B when no control is imposed. If constraining an agent's choice set not only lowers the intrinsic motivation to behave pro-socially towards principal B whose transfer is affected by the constraint, but lowers the agent's motivation to behave pro-socially towards *any* principal, transfers to C may also be lower when control is imposed. Similarly, if agents interpret the control decision as an indication about the social norm for agents' behavior towards principals in general, transfers to C may be negatively affected by the control decision. Both mechanisms may lead to detrimental indirect effects of control in the IN as well as the EX treatments to the same extent. If, however, the negative reaction to control hinges on the direct relation between the constrained agent and the controller's earnings, the transfer to C may not be affected by the control decision.

We now consider the behavior of selfish individuals, that would transfer less than the minimum transfer to principal B if no control was imposed. Note, that in our IN treatments an action asymmetry between the two principals exist, as only one of them makes an active decision about whether to constrain the agent in the transfer to him or not. In the EX treatments instead, this asymmetry does not exist. Agents may have a preference to treat the two principals equally if they have equal opportunities in the sense of not having a choice option. In that case when selfish agents have to increase their transfer to principal B when externally controlled, they may adjust the transfer to principal C accordingly. In the IN treatments instead, due to the choice inability of principal C, no such symmetry between the principals exists, so that positive spill-over effects may be less likely to occur.

5.6 Results

We first look at the agents' behavioral reaction to control and investigate whether direct and indirect detrimental effects of control exist. We then evaluate the impact of those effects on overall transfers and take differences in agents' perception of control into account in our analysis.

5.6.1 Behavioral reactions to control

We first look at how agents' transfers to principal B are cumulatively distributed given the decision of the controller in the four treatments. Following the analysis in FK we reason that if the constraint did not have any effect on the agent's transfer to B, at any $x_B \geq \underline{x}_B$ the cumulative distributions of agents' choices in the condition that the controller constrains would coincide with the cumulative distributions in case the controller does not constrain. Figure 5.2 shows the cumulative distribution functions for all four treatments. In IN10 for each value of $x_B \geq \underline{x}_B$, there are always more or equally many agents in the no constraint condition who choose x_B such that it is at least that value in the constraint condition. In IN5 the relation between the two cumulative distribution functions is less systematic. In order to test whether the shifts in the distributions are statistically significant, as proposed by FK, we replace all $x_B < \underline{x}_B$ by $\underline{x}_B$ in the no constraint condition. If the constraint did not have any impact the resulting modified distribution should not be different from that in the constraint condition. A Wilcoxon signed-ranks (WS) test yields that in the IN10 treatment the two distributions are significantly different (p-value=0.0471): the cumulative distribution function in the unconstrained condition is always below that in the constrained condition.[7] In the IN5 treatment, the two distributions are not significantly different from each other (p-value=0.2033). As we only find significant evidence for hidden costs of control in the IN10 treatment but not in the IN5 treatment, we interpret our findings as weak evidence for hidden costs of control. The results are surprising as, in contrast to our findings, FK find that the lower the constraint, the higher the hidden costs of control.

[7] The applied tests are always 2-sided.

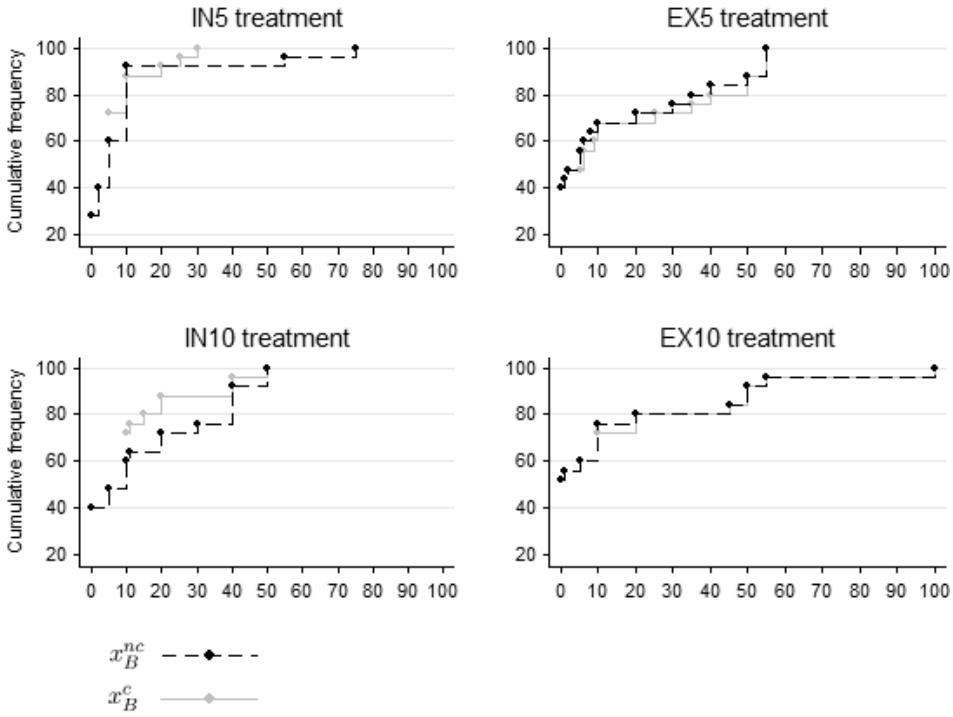


Fig. 5.2: Cumulative frequencies of transfers to principal B

In EX5 and EX10 for each value of $x_B \geq \underline{x}_B$, there are always equally many or more agents in the constrained condition who choose x_B such that it is at least that value in the unconstrained condition. The differences are however statistically insignificant for both treatments (WS test, EX5: p=0.1154; EX10: p=0.3173). The corresponding figure for the cumulative distributions of x_C^{nc} and x_C^c can be found in Appendix 5.8.1. As for the analysis of the uncontrollable transfer to C no distribution needs to be truncated, the results of the statistical tests are shifted to Section 5.6.2 in which average transfers are discussed.

Result. *We find evidence for a direct negative effect of control in the transfers to principal B in the IN10 treatment but not in the IN5, EX5 and EX10 treatment.*

Table 5.1 shows the number and relative share of individuals in aggre-

gation for the two IN and the two EX treatments respectively that show a positive, neutral or negative *direct* reaction to control via the transfer to principal B. Conditional on this direct reaction the table also shows the number and relative shares of individuals that have a positive, neutral or negative *indirect* reaction to control via the transfer to principal C.
For a majority of individuals their *direct* reaction to control via the transfer to principal B is positive irrespective of whether the controller is internal or external (48% and 62% respectively). So clearly, control also generates benefits as selfish agents are forced to transfer at least $\underline{x}_B$ when constrained. Whereas there is only one individual out of 50 whose direct reaction to control is negative in the EX treatments (2%), in the IN treatments the fraction of individuals that react negatively is considerable (26%).

Table 5.1: Direct and indirect behavioral reactions to control

				IN treatments					
x_B		positive	(+)		neutral	(=)		negative	(-)
Number of agents		24			13			13	
Relative share		48%			26%			26%	
x_C [conditional on x_B]	+	=	-	+	=	-	+	=	-
Number of agents	4	16	4	0	12	1	4	6	3
Relative share	17%	67%	17%	0%	92%	8%	31%	46%	23%

				EX treatments					
x_B		positive	(+)		neutral	(=)		negative	(-)
Number of agents		31			18			1	
Relative share		62%			36%			2%	
x_C [conditional on x_B]	+	=	-	+	=	-	+	=	-
Number of agents	11	19	1	0	18	0	1	0	0
Relative share	36%	61%	3%	0%	100%	0%	100%	0%	0%

Result. *We observe a strong heterogeneity in agents' direct reaction to the controller's implementation of control: they react negatively, neutrally or positively. The latter group is always the majority. If the controller is external the fraction of agents with a negative direct reaction is negligible, whereas it is considerable when control is internal.*

We now turn to the indirect effect of control via the transfer to principal C. We find that for a majority of individuals the *indirect* reaction to control is neutral in both, the IN and EX treatments (IN: 65%; EX: 74%). Given a certain direct reaction to control via the transfer to principal B we do not find any significant difference between x_C^{nc} and x_C^{c} when the controller is internal (WS-test, p-values≥0.3173). We conclude that agents with a negative (positive) direct reaction to control do not systematically react negatively (positively) to control via the transfer to principal C.
With an external controller we find that a considerable fraction of agents (36%) that show a positive direct reaction to being controlled, also show an indirect positive reaction. The difference between x_C^{c} and x_C^{nc} among those agents is statistically significant with higher transfers to principal C when control is imposed (p-value=0.0029).
All agents with a neutral direct reaction to control in the EX treatments also have an indirect neutral reaction to control.

Result. *For a majority of agents the indirect reaction to control is neutral. When the controller is internal we do not find evidence for systematic spillover effects to the transfer to C. If the controller is external instead, we find that a considerable fraction of agents with a positive direct reaction to control also transfer more to principal C when controlled.*

5.6.2 Average transfers

Individuals' behavior translates into the average transfers depicted in Figure 5.3. Note, that the maximum transfer per principal is 100 points and that a transfer of 55 points to each of the two principals would equalize participants' earnings.
Transfers to principal B *and* to principal C in the IN treatments are depicted in the two graphs on the left, while those for the EX treatments are

on the right. In the IN5 treatment transfers to both principals are not significantly different when control is imposed versus not imposed (WS-test, p-values≥0.4397). The same is true for the IN10 treatment (WS-test, p-values≥0.4969). In contrast, in the EX5 and the EX10 treatment transfers to B *and* to C are significantly higher when there is control as compared to when there is no control (WS-test, p-values≤0.0085).

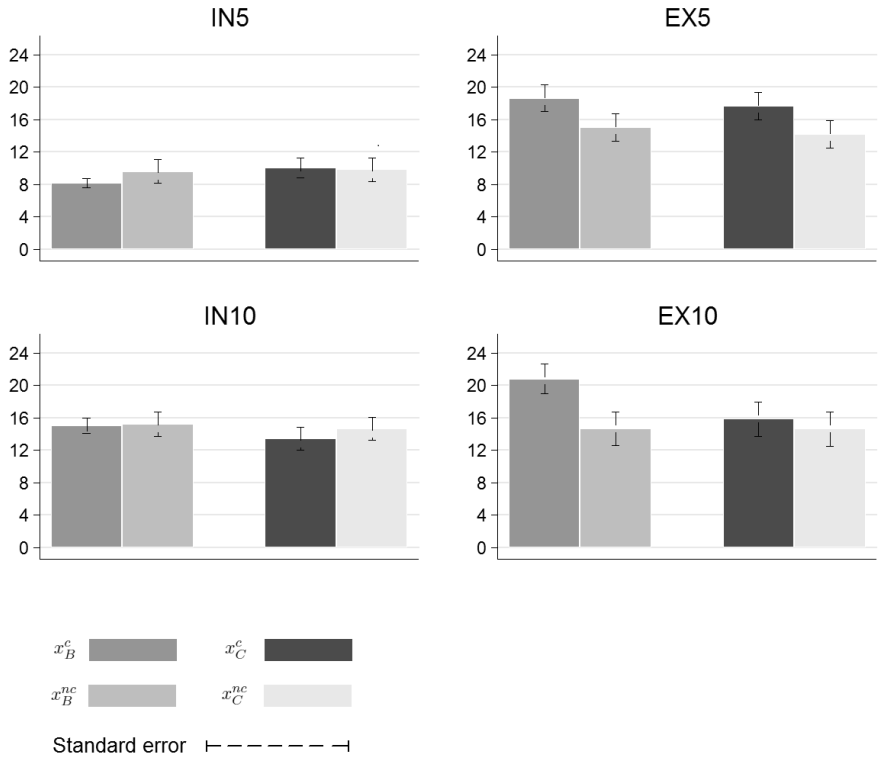


Fig. 5.3: Average transfer to principal B and principal C

Result. *Irrespective of the level of the minimum transfer to B, control does not have an effect on average transfers when the controller is internal. When he is external instead, agents transfer significantly more to B and to C when control is imposed as compared to when the controller refrains from imposing it.*

As a majority of 64% of the controllers in the IN treatments and 88% in the EX treatments in fact decide to impose control, agents' transfer decisions in the control condition are most relevant for the eventual distribution of earnings between principals and agents. For a low constraint of 5 a Mann Whitney (MW) test yields that transfers x_B^c in EX5 are significantly higher than in IN5 (p-value=0.0535). For transfers x_C^c this difference is insignificant (p-value=0.2219). With a high constraint of 10 transfers x_B^c in EX10 are not significantly different from those in IN10 and so are transfers x_C^c (p-value≥0.7222). Comparing x_C^{nc} and x_B^{nc} between treatments for either of the two constraint levels yields that differences are statistically insignificant (MW-test, p-values≥0.3983).

5.6.3 Agents' perception of control

The differences in the occurrence of negative direct reactions to control between the IN treatments and the EX treatments may occur for two reasons. First, negative reactions may only occur in the former case because of the direct relation between the controller and the agent in that setting: the constraint decision is made by a principal whose earnings are at stake, and the agent has the possibility to directly react towards the controller with his effort provision. A second reason may be that the imposition of internal control is perceived differently than that of external control, which then results in different effects on the agents' motivation. FK finds that most agents who react negatively to being controlled perceive the controller's constraint decision as a signal of distrust or a limitation of their choice autonomy. We thus want to compare individuals' perception of control between the IN and the EX treatments with respect to i.a. these factors. Participants were asked to answer a set of questions that elicited to what extent they agree or disagree to perceive being constrained as controlling, distrusting, a lack of freedom and whether it is okay for them to be constrained (see questions 1-4 in Appendix 5.8.4).

Separating the results provided in Table 5.1 by the constraint level, we observe that 25% of the agents in the IN10 treatment (6 out of 25) react negatively to control, whereas nobody does so in the EX10 treatment. A MW test yields that no significant differences in the agreement to the four perception

statements exist between IN10 and EX10 (p-values≥0.4649). This suggests that the differences in the agents' behavioral reaction to control when it is internal as compared to external are not driven by differences in how they perceive control with respect to the underlying factors.

In the IN5 treatment 28% (7 out of 25) react negatively to control, whereas only 4% (1 out of 25) do so in EX5. Comparing agents' perception between the two treatments we find that they agree equally much to perceive the constraint as controlling and as a lack of freedom, and to be okay with being constrained (p-values≥ 0.4350). However, we find a difference in their perception of distrust. In IN5 agents agree significantly more that they feel distrusted when they are constrained as compared to EX5 (p-value=0.0124).

We further find that among those agents that can fall into the 'negative reaction to control' group (agents with $x_B^{nc} > \underline{x}_B$), reciprocity attitudes and justice perceptions are not different between those that indeed react negatively and those that do not (p-values≥ 0.3602).

5.7 Discussion and conclusion

We experimentally investigate how control affects agents' behavior in a multiple principals-agent game when it is either imposed by a controller whose earnings are not at stake (external control) or by one whose earnings are determined by the agent's effort provision (internal control). Further, we investigate whether an agent's direct behavioral reaction to control spills over to his relation with another principal in which control can not be imposed.

Agents are asked to decide how much they are willing to transfer to two principals B and C in one of two condition: a) the choice set for the transfer to B is constrained to a certain minimum transfer and b) the constraint is not imposed and the agent is free to transfer any amount to both of them. The transfer to principal C can never be constrained.

We find that when control is internal a considerable fraction of agents reduce their effort towards the controlling principal B as compared to when he does not control. Their uncontrollable effort provided to principal C is not systematically affected by the control decision. We do not find significant differences in average transfers to both principals between the constrained and the unconstrained condition. In contrast, when control is external, ef-

fort provision is significantly higher for both principals when control is imposed as compared to when there is no control. For a low minimum effort requirement we find that the agents' transfers to principal B under control are significantly higher when control is imposed by an external controller as compared to an internal one.

Controlling for differences in the perception of control in our analysis, our results suggest that the occurrence of detrimental effects crucially depends on who is the controller, i.e. whether it is the principal who profits (suffers) from high (low) effort provision by the agent or an external entity whose earnings are not at stake. Only in the case of an internal controller we find that several agents react negatively to control. This suggests that the driving mechanism behind detrimental effects of control is not the interaction with the agents' intrinsic motivation to behave pro-socially towards the principals or the signal of the control decision about the social norm of effort provision which is then followed. Instead, it seems that hidden costs of control work via a mechanism in which control provides information about the controller and/or his action to which the agents react or reciprocate via their effort provision.

We conclude that in the aggregate control can be effective in enforcing high effort provision if a third party, e.g. the government or a consultancy, makes the control decision. This result offers a rationale for the delegation of the control decision to an external entity. Further research, however, is necessary to investigate whether the responsibility for the control decision can effectively be shifted away from the internal principal if he delegates the decision to an external controller.

5.8 Appendix

5.8.1 Additional figure

Figure 5.4 shows how agents' transfers x_C are cumulatively distributed given the decision of the principal in the four treatments.

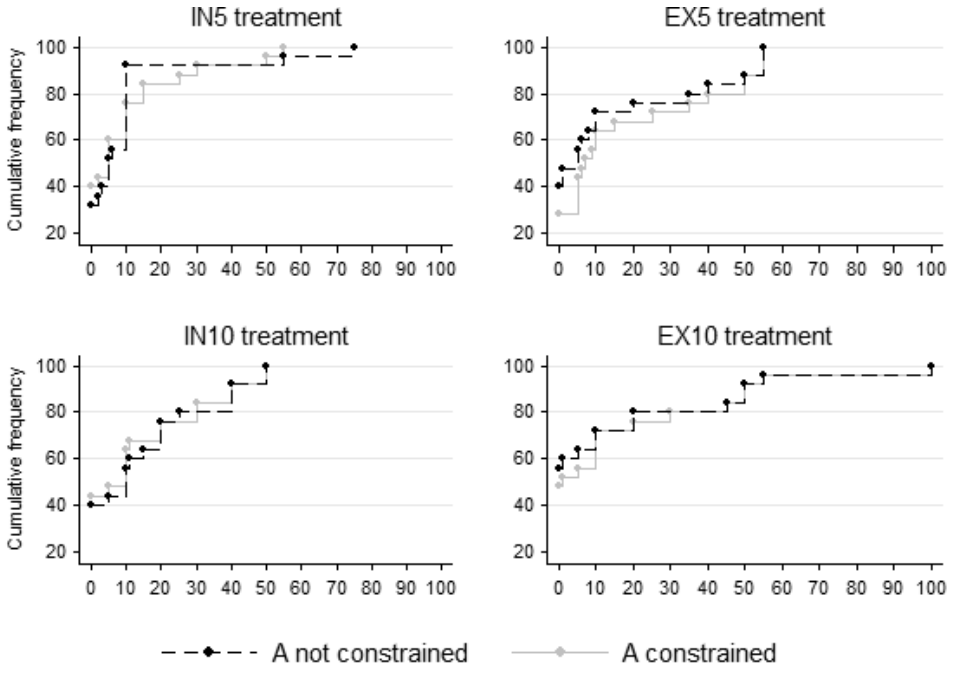


Fig. 5.4: Cumulative frequencies of transfers to principal C

In EX5 and EX10 for each value of x_C, there are always equally many or more agents in the constraint condition than in the no constraint condition who choose a certain x_C.

5.8.2 Decision screens

Below you can find decision screens for the controller (Figure 5.5) and the agent (Figure 5.6 and Figure 5.7), exemplary for treatment IN5.

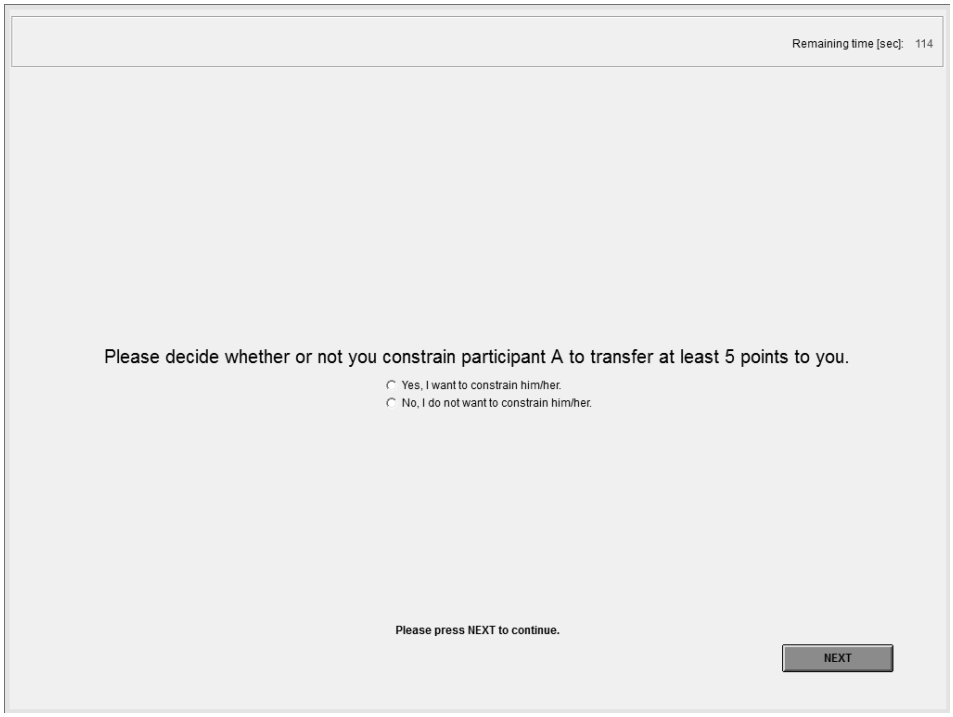


Fig. 5.5: Decision screen controller

Fig. 5.6: Decision screen agent I

Fig. 5.7: Decision screen agent II

5.8.3 Experimental instructions

[These are exemplary instructions for $\underline{x}_B = 5$.]

Thank you for participating in this decision-making experiment. **Please read the following instructions carefully!**

If there is something that you do not understand please ask for assistance. Do not speak out loud but raise your hand. An experimenter will then come to you and answer your question. During the experiment there is a strict prohibition of communication other than described in these instructions.

You receive 3 euros for having shown up at the experiment in time. You can earn additional money in the course of the experiment by collecting points. All points you earn will be converted to euros at the end of the experiment, using the exchange rate of 1 point = 8 eurocents.

At the end of the experiment, you will be paid out privately and in cash what you have earned in the course of the experiment plus the mentioned fixed payment of 3 euros. Your final payment will be rounded in increments of 5 eurocents.

Specific instructions

In this experiment, each participant is randomly matched with two other participants to form a group of three. The participants in the group are called A, B, and C. Neither during nor after the experiment will anybody get to know the identities of the other matched participants. Hence, all decisions remain anonymous.

After you have finished reading these instructions, you will be informed on the computer screen in front of you whether you are a participant A, B or C.

Initial budget: Participant A receives an initial budget of 220 points.
Participant B receives an initial budget of 0 points.
Participant C receives an initial budget of 0 points.

Participant A:

Participant A has to decide how many points, if any, s/he wants to transfer

to participant B and how many points s/he wants to transfer to participant C. Each point transferred to B and/or C will be doubled. That is, each point A transfers to B reduces A's income by one point and increases B's income by two points. Likewise, each point A transfers to C reduces A's income by one point and increases C's income by two points.

The formulas for calculating income look as follows:
Participant A's income: 220 - transfer to B - transfer to C
Participant B's income: 0 + 2*transfer from A to B
Participant C's income: 0 + 2*transfer from A to C

Participant A can transfer any amount between 0 and 100 points to participant B. Likewise, Participant A can also transfer any amount between 0 and 100 to participant C. There may be further constraints on the amount that participant A can transfer to B, which are described below.

[From here onwards the instructions differed between treatment IN and EX]

Treatment EX

Participant B* of another group:

Before participant A decides how many points s/he wishes to transfer to participant B and participant C, a randomly selected participant B of another group, henceforth called B* of another group, has the possibility to constrain participant A to transfer at least 5 points to participant B. Participant B* of another group can also decide not to constrain participant A and thus leave her/his transfer decision completely free. Importantly, the income of B* of another group does not depend in any way on the decision of participant A, but is in turn determined by the decision of participant A* in his/her own group. Note that this participant A* of the other group may be constrained by a participant B of yet another group. Irrespective of whether B* of another group constrains participant A or not, A can always transfer any amount between 0 and 100 to participant C.

Therefore, there are **two cases**...

... Participant B* of *another* group constrains participant A to transfer at least 5 points to participant B. In this case, participant A can transfer any

(integer) amount between 5 and 100 to B, and any (integer) amount between 0 and 100 to C.

... Participant B* of *another* group does not constrain participant A to transfer at least 5 points to participant B. That is, s/he allows participant A to decide on her/his transfer freely. In this case, participant A can transfer any (integer) amount between 0 and 100 to B and any (integer) amount between 0 and 100 to C.

Participant B (of your group):

Participant B (of your group) has the possibility to constrain participant A of *yet another randomly chosen* group to transfer at least 5 points to participant B in that other group. Participant B can also decide not to constrain participant A of that other group and thus leave her/his transfer decision completely free. Neither participant A nor C (of your group) will be informed about the decision B makes regarding that other group. Irrespective of whether B constrains A of that other group or not, A in that other group can always transfer any amount between 0 and 100 to participant C in that other group.

Participant C:

Participant C does not make a decision.

In summary, the experiment consists of two stages:

Stage 1: In stage 1, participant B* of *another* group decides if s/he constrains participant A to transfer at least 5 points to participant B, or if s/he allows participant A to decide freely.
Participant B decides if s/he constrains participant A of *yet another* group to transfer at least 5 points to participant B of that other group, or if s/he allows that participant A to decide freely.

Stage 2: In stage 2, participant A decides on the amount s/he will transfer to participant B and to participant C. Please note that participant A will have to decide before s/he knows whether or not participant B* of *another* group constrains her/him. That is, participant A has to indicate her/his transfer decisions for **both possible cases that**...

... Participant B* of *another* group constrains participant A to transfer at

least 5 points to participant B. The amount A transfers to B can be at least 5 points and at most 100 points and the amount A transfers to C can be at least 0 points and at most 100 points.

... Participant B* of *another* group does not constrain participant A in the transfer to participant B. The amount A transfers to B can be at least 0 points and at most 100 points and the amount A transfers to C can be at least 0 points and at most 100 points.

The actual decision of B* of another group determines which of the two cases is relevant for the incomes. If participant B* of another group constrains participant A to transfer at least 5 points to participant B, the transfer decisions which A indicates for this particular case apply. If participant B* of another group leaves participant A free in his/her decision, the transfer decisions which participant A indicates for that case apply.

Note: When A makes her/his decision s/he does not know the decision of B* of another group. Therefore, A should consider each of the two cases as the case that is relevant for the incomes.

B's and C's incomes are determined solely by the transfer decisions of participant A.

It is important that all participants have correctly understood the instructions and income rules. Thus, you will now take part in a practice round. This practice round will not be relevant for your income. Please activate the computer screen in front of you by moving the mouse and further instructions will follow on-screen.
At the end of the experiment a final income screen will inform you about the decisions that were made and the resulting incomes in your group. As already said at the beginning of the instructions, your income in points will be converted to euros and together with the fixed payment of 3 euros paid out in cash to you in private.

Treatment IN

Participant B:

Before participant A decides how many points s/he wishes to transfer to participant B and participant C, participant B has the possibility to constrain participant A to transfer at least 5 points to her/him. B can also decide not to constrain participant A and thus leave her/his transfer decision completely free. Irrespective of whether B constrains A or not, A can always transfer any amount between 0 and 100 to C. Therefore, there are **two cases**...

... Participant B constrains participant A to transfer at least 5 points to her/him. In that case, participant A can transfer any (integer) amount between 5 and 100 to B, and any (integer) amount between 0 and 100 to C.
... Participant B does not constrain participant A to transfer at least 5 points to her/him. That is, s/he allows participant A to decide on her/his transfer freely. In this case, participant A can transfer any (integer) amount between 0 and 100 to B, and any (integer) amount between 0 and 100 to C.

Participant C:

Participant C does not make a decision.

In summary, the experiment consists of two stages:

Stage 1: In stage 1, participant B decides if s/he constrains participant A to transfer at least 5 points to her/him (participant B), or if s/he allows participant A to decide freely.

Stage 2: In stage 2, participant A decides on the amount s/he will transfer to participant B and to participant C. Please note that participant A will have to decide before s/he knows whether or not participant B constrains her/him. That is, participant A has to indicate her/his transfer decisions for **both possible cases that**...

... Participant B constrains participant A to transfer at least 5 points to her/him. The amount A transfers to B can be at least 5 points and at most 100 points and the amount A transfers to C can be at least 0 points and at most 100 points.

... Participant B does not constrain participant A in the transfer to her/him. The amount A transfers to B can be at least 0 points and at most 100 points and the amount A transfers to C can be at least 0 points and most 100 points.

The actual decision of B determines which of the two cases is relevant for the incomes. If participant B constrains participant A to transfer at least 5 points to her/him, the transfer decisions which A indicates for this particular case apply. If participant B leaves A free in his/her decision, the transfer decisions which participant A indicates for that case apply.

Note: When A makes her/his decision s/he does not know the decision of B. Therefore, A should consider each of the two cases as the case that is relevant for the incomes.

B's and C's incomes are determined solely by the transfer decisions of participant A.

It is important that all participants have correctly understood the instructions and income rules. Thus, you will now take part in a practice round. This practice round will not be relevant for your income. Please activate the computer screen in front of you by moving the mouse and further instructions will follow on-screen.
At the end of the experiment a final income screen will inform you about the decisions that were made and the resulting incomes in your group. As already said at the beginning of the instructions, your income in points will be converted to euros and together with the fixed payment of 3 euros paid out in cash to you in private.

5.8.4 Questionnaires

Questionnaire 'Constraint'

In this questionnaire the agents are asked to indicate to what extent they personally agrees or disagree to the following statements [1 = "strongly disagree", ..., 4 = "neutral", .. ., 7 = "strongly agree"].

EX5 treatment

- I perceive the constraint by participant B* of another group to transfer at least 5 points to participant B as controlling.

- I feel distrusted when constrained by participant B* of another group to transfer at least 5 points to participant B.
- I perceive it as a lack of freedom of choice when constrained by participant B* of another group to transfer at least 5 points to participant B.
- I find it okay that participant B* of another group constrains me to transfer at least 5 points to participant B.
- I feel that a person who constrains me to transfer at least 5 points to participant B is a greedy person.
- I do not have any particular opinion about being constrained by participant B* of another group to transfer at least 5 points to participant B.
- If participant B (of my group) had had the choice to constrain me to transfer at least 5 points to him/her, s/he would have done it.
- If participant C (of my group) had had the choice to constrain me to transfer at least 5 points to him/her, s/he would have done it.

IN5 treatment

- I perceive the constraint by participant B to transfer at least 5 points to participant B as controlling.
- I feel distrusted when constrained by participant B to transfer at least 5 points to participant B.
- I perceive it as a lack of freedom of choice when constrained by participant B to transfer at least 5 points to participant B.
- I find it okay that participant B constrains me to transfer at least 5 points to participant B.
- I feel that a person who constrains me to transfer at least 5 points to him/her is a greedy person.

- I do not have any particular opinion about being constrained by participant B to transfer at least 5 points to participant B.
- If participant C had had the choice to constrain me to transfer at least 5 points to him/her, s/he would have done it.

Questionnaire 'Reciprocity'

Respondents were asked to indicate on a 7-point scale how well each of the following six statements applies to them personally:

- If someone does me a favour, I am prepared to return it;
- If I suffer a serious wrong, I will take revenge as soon as possible, no matter what the cost;
- If somebody puts me in a difficult position, I will do the same to him/her;
- I go out of my way to help somebody who has been kind to me before;
- If somebody offends me, I will offend him/her back;
- I am ready to undergo personal costs to help somebody who helped me before.

An answer of 1 on the scale means: 'does not apply to me at all' and choosing 7 means: 'applies to me perfectly'. Questions (1), (4) and (6) ask about positive reciprocity, while questions (2), (3) and (5) ask about negative reciprocity. Also, two of the questions ask explicitly whether the respondent would incur costs in order to be negatively reciprocal (question 2) or positively reciprocal (question 6).

Questionnaire 'Just world'

Participants were asked to decide to what extent they personally agree or disagree with the following statements.

- I think basically the world is a just place.
- I believe that, by and large, people get what they deserve.

- I am confident that justice always prevails over injustice.
- I am convinced that in the long run people will be compensated for injustices.
- I firmly believe that injustices in all areas of life (e.g., professional, family, politic) are the exception rather than the rule.
- I think people try to be fair when making important decisions.
- I believe that, by and large, I deserve what happens to me.
- I am usually treated fairly.
- I believe that I usually get what I deserve.
- Overall, events in my life are just.
- In my life injustice is the exception rather than the rule.
- I believe that most of the things that happen in my life are fair.
- I think that important decisions that are made concerning me are usually just.

An answer of 1 on the scale means: 'strongly disagree' and choosing 6 means: 'strongly agree.

Bibliography

Alesina, A. and Ferrara, E. L. (2005). Preferences for redistribution in the land of opportunities. *Journal of Public Economics*, 89:897–931.

Alesina, A. and Glaeser, E. L. (2004). Fighting poverty in the US and Europe: a world of difference. *Oxford University Press, Oxford.*

Attanasio, O., Barr, A., Cardenas, J. C., Genicot, G., and Meghir, C. (2012). Risk pooling, risk preferences, and social networks. *American Economic Journal: Applied Economics*, 4(2):134–167.

Bardsley, N. (2008). Dictator game giving: altruism or artifact? *Experimental Economics*, 11:122–133.

Barr, A. and Genicot, G. (2008). Risk sharing, commitment and information: An experimental analysis. *Journal of the European Economic Association*, 6(6):1151–1185.

Bartling, B. and Fischbacher, U. (2012). Shifting the blame: On delegation and responsibility. *The Review of Economic Studies*, 79(1):67–87.

Bellemare, C., Kroöger, S., and van Soest, A. (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica*, 76 (4):815–839.

Belot, M. and Schröder, M. (2013). Does monitoring work? a field experiment with multiple forms of counterproductive behaviour. *Working paper No. 130006, Otto-von-Guericke University Magdeburg,*.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*, 57:289–300.

Bolton, G. and Ockenfels, A. (2000). A theory of equity, reciprocity and competition. *American Economic Review*, 100:166–193.

Bolton, G. E. and Katok, E. (1995). An experimental test for gender differences in beneficent behavior. *Economics letters*, 48:287–292.

Boly, A. (2011). On the incentive effects of monitoring: evidence from the lab and the field. *Experimental Economics*, 14(2):241–253.

Bossert, W. and Fleurbaey, M. (1996). Redistribution and compensation. *Social Choice and Welfare*, 13(3):342–355.

Bowles, S. and Polania-Reyes, S. (2012). Economic incentives and social preferences: substitutes or complements? *Journal of Economic Literature*, 50(2):368–425.

Cabrales, A., Nagel, R., and Mora, J. V. R. (2006). It is hobbes, not rousseau: an experiment on social insurance. *UC3M Working Paper 07-08*.

Camerer, C. F. (2003). Behavioral game theory: experiments in strategic interaction. *Princeton: Princeton University Press*.

Cappelen, A. W., Hole, A. D., Sorensen, E., and Tungodden, B. (2007). The pluralism of fairness ideals: An experimental approach. *American Economic Review*, 97(3):818–827.

Cappelen, A. W., Konow, J., Sorensen, E., and Tungodden, B. (2013). Just luck: An experimental study of risk taking and fairness. *American Economic Review*, 103(4):1398–1413.

Carpenter, J., Verhoogen, E., and Burks, S. (2005). The effect of stakes in distribution experiments. *Economic Letters*, pages 393–398.

Cawley, J. and Ruhm, C. (2011). The economics of risky health behaviors. *NBER Working Paper 17081*.

Centraal Planbureau (2000). Solidariteit, keuzevrijheid, en transparantie. de toekomst van de nederlandse markt voor oudedagsvoorzieningen (solidarity, freedom of choice, and transparency. the future of the market for old-age security in the netherlands). *Den Haag, SDU Uitgevers*.

Cettolin, E. and Tausch, F. (2013). Risk sharing and risk taking: does responsibility matter? *UM Working Paper RM/13/045*.

Charness, G. and Genicot, G. (2009). Informal risk sharing in an infinite-horizon experiment. *The Economic Journal*, 119(537):796–825.

Charness, G. and Gneezy, U. (2008). What's in a name? anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior and Organization*, 68(1):29–35.

Chaudhuri, A., Gangadharan, L., and Maitra, P. (2010). An experimental analysis of group size and risk sharing. *Working paper. Available online at http://users.monash.edu.au/~maitra/Risksharing_march26_2010.pdf*.

Chavanne, D., McCabe, K., and Paganelli, M. P. (2009). Redistributive justice–entitlements and inequality in a third-party dictator game. *George Mason University Working Paper*.

CNN on the web (2011). Should smokers and fat people pay more for health care?. *Cafferty, J.*, October 31.

Coffman, L. C. (2011). Intermediation reduces punishment (and reward). *American Economic Journal: Microeconomics*, 3(4):77–106.

Cohen, A. and Siegelman, P. (2010). Testing for adverse selection in insurance markets. *Journal of Risk and Insurance, The American Risk and Insurance Association*, 77(1):39–84.

Cox, J. C. and Deck, C. A. (2006). When are women more generous than men? *Economic Inquiry*, 44(4):587–598.

Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic literature*, 47(2):448–474.

Cutler, D. M., Finkelstein, A., and McGarry, K. (2008). Preference heterogeneity and inmarkets markets: explaining a puzzle of insurance markets. *NBER Working Paper 13746*.

Dalbert, C. (2000). *Beliefs in a just worls questionnaire.* In J. Maltby, C.A. Lewis, & A. Hill (Eds.), Commissioned reviews of 250 psychological Tests (pp. 461-465), Lampeter, Wales: Edwin Mellen Press.

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology*, 18(1):105.

D'Excelle, B. and Riedl, A. (2010). Directed generosity and network formation: network dimension matters. *IZA Discussion Paper No. 5356*.

Dickinson, D. L. and Tiefenthaler, J. (2002). What is fair? Experimental evidence. *Southern Economic Journal*, pages 414–428.

Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2009). Homo reciprocans: Survey evidence on behavioural outcomes*. *The Economic Journal*, 119(536):592–612.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Durante, R. and Putterman, L. (2009). Preferences for redistribution and perception of fairness: An experimental study. *Mimeo. Brown University*.

Eckel, C. C. and Grossman, P. J. (1998). Are women less selfish than men?: Evidence from dictator experiments. *The economic journal*, 108(448):726–735.

Einav, L., Finkelstein, A., Pascu, I., and Cullen, M. R. (2012). How general are risk preferences? choices under uncertainty in different domains. *American Economic Review*, 102(6):2606–38.

Ellingsen, T. and Johannesson, M. (2008). Pride and prejudice: The human side of incentive theory. *American Economic Review*, 98(3):990–1008.

Esarey, J., Salmon, T., and Barrilleaux, C. (2009). Social insurance and income redistribution in a laboratory experiment. *available at SSRN: hhtp://ssrn.com/abstract=1361134*.

Falk, A. and Kosfeld, M. (2006). The hidden costs of control. *American Economic Review*, 96(5):1611–1630.

Faravelli, M. (2007). How context matters: A survey based experiment on distributive justice. *Journal of Public Economics*, 91(7-8):1399–1422.

Fehr, E. and Falk, A. (1999). Wage rigidity in a competitive incomplete contract market. *Journal of political Economy*, 107(1):106–134.

Fehr, E. and Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature*, 422(6928):137–140.

Fehr, E. and Schmidt, K. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics*, 114:817–868.

Fenger, M. (2009). Challenging solidarity? an analysis of exit options in social policies. *Social Policy & Administration*, 43(6):649–665.

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

Fong, C. (2011). Social preferences, self-interest and the demand for redistribution. *Journal of Public Economics*, 822:225–246.

Forsythe, R. J., Horowitz, J., Savin, N., and Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3):347–369.

Fowler, J. and Kam, D. (2007). Beyond the self: Social identity, altruism, and political participation. *Journal of Politics*, 69(3):811–825.

Fréchette, G. R. (2011). Laboratory experiments: Professionals versus students. *Available online at http://ssrn.com/abstract=1939219*.

Frey, B. S. and Benz, M. (2001). Motivation transfer effect. *manuscript*.

Frey, B. S. and Jegen, R. (2001). Motivation crowding theory. *Journal of economic surveys*, 15(5):589–611.

Frohlich, N. and Oppenheimer, J. A. (1990). Choosing justice in experimental democracies with production. *The American Political Science Review*, pages 461–477.

Frohlich, N., Oppenheimer, J. A., and Eavey, C. L. (1987). Choices of principles of distributive justice in experimental groups. *American Journal of Political Science*, pages 606–636.

Gächter, S. and Riedl, A. (2005). Moral property rights in bargaining with infeasible claims. *Management Science*, 51(2):249–263.

Gächter, S. and Riedl, A. (2006). Dividing justly in bargaining problems with claims. *Social Choice and Welfare*, 27(3):571–594.

Gneezy, U., Meier, S., and Rey-Biel, P. (2011). When and why incentives (don't) work to modify behavior. *The Journal of Economic Perspectives*, 25(4):191–209.

Gneezy, U. and Rustichini, A. (2000). Pay enough or don't pay at all. *The Quarterly Journal of Economics*, 115(3):791–810.

Goeree, J., McConnell, M., Mitchell, T., Tromp, T., and Yarif, L. (2010). The 1/d law of giving. *american economic journal:micro economics*, 2(1):183–203.

Greiner, B. (2004). An online recruitment system for economic experiments. *In: Kurt Kremer, Volker Macho (Hrsg.): Forschung und wissenschaftliches Rechnen. GWDG Bericht 63. Ges. für Wiss. Datenverarbeitung, Göttingen*, pages 79–93.

Güth, W., Offerman, T., Potters, J., Strobel, M., and Verbon, H. A. A. (2002). Are family transfers crowded out by public tramsfers? *Scandinavian Journal of Economics*, 104(4):587–604.

Hagemann, P. (2007). What's in a frame? on demand effects and trust in experimenta studies. *manuscript*.

Hammond, P. (1975). Charity: Altruism or cooperative egoism. *In Phelps, E. (ed), Altruism, morality and economic theory. New York: Sage*, pages 115–131.

Harrison, G. and Cox, J. (2008). *Risk Aversion in Experiments*. Research in Experimental Economics Series. Emerald Books.

Harsanyi, J. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparison of well-being. *Journal of Political Economy*, 63:309–321.

Herne, K. and Suojanen, M. (2004). The role of information in choices over income distributions. *Journal of Conflict Resolution*, 48(2):173–193.

Höchtl, W., Sausgruber, R., and Tyran, J.-R. (2012). Inequality aversion and voting on redistribution. *European economic review*, 56(7):1406–1421.

Hoffman, E., McCabe, K., Shachat, K., and Smith, V. L. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, 7(3):346–380.

Hoffman, E., McCabe, K., and Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *The American Economic Review*, 86(3):653–660.

Jakiela, P. (2009). How fair shares compare: experimental evidence from two cultures. *Mimeo*.

Johannesson, M. and Persson, B. (2000). Non-reciprocal altruism in dictator games. *Economics Letters*, 69(2):137–142.

Kessler, J. B. and Leider, S. (2013). Finding the cost of control. *CESifo working paper no. 4188*.

Klor, E. and Shayo, M. (2010). Social identity and preferences over redistribution. *Journal of Public Economics*, 94(3):269–278.

Konow, J. (1996). A positive theory of economic fairness. *Journal of Economic Behavior & Organization*, 31(1):13–35.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review*, 90(4):1072–1091.

Konow, J., Saijo, T., and Akai, K. (2009). Morals and mores? experimental evidence on equity and equality. *Mimeo*.

Kotlikoff, L. J., Persson, T., and Svensson, L. E. O. (1988). Social contracts as assets: A possible solution to the time-consistency problem. *The American Economic Review*, 78(4):662–677.

Krawczyk, M. (2010). A glimpse through the veil of ignorance: Equality of opportunity and support for redistribution. *Journal of Public Economics*, 94:131–141.

Krawczyk, M. and Le Lec, F. (2008). Social decisions under risk. evidence from the probabilistic dictator game. *Mimeo*.

Leider, S., Mobius, M., Rosenblat, T., and Do, Q. (2009). Directed altruism and enforced reciprocity in social networks. *Quarterly Journal of Economics*, 124(4):1815–1851.

List, J. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy*, 115(3):482–493.

List, J. A. and Cherry, T. L. (2008). Examining the role of fairness in high stakes allocation decisions. *Journal of Economic Behavior & Organization*, 65(1):1–8.

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

Messer, K. D., Poe, G. L., Rondeau, D., Schulze, W. D., and Vossler, C. A. (2010). Social preferences and voting: An exploration using a novel preference revealing mechanism. *Journal of Public Economics*, 94(3):308–317.

New York Times on the web (2011). The smokers' charge. *Abelson R.*, November 17.

Normann, H.-T. and Ricciuti, R. (2009). Laboratory experiments for economic policy making. *Journal of Economic Surveys*, 23(3):407–432.

Offerman, T., Potters, J., and Verbon, H. A. (2001). Cooperation in an overlapping generations experiment. *Games and Economic Behavior*, 36(2):264–275.

Oxoby, R. J. and Spraggon, J. (2008). Mine and yours: Property rights in dictator games. *Journal of Economic Behavior & Organization*, 65(3):703–713.

Paetzel, F., Sausgruber, R., and Traub, S. (2012). Social preferences and voting on reform-an experimental study. *Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2066455*.

Rawls, J. (1971). *A Theory of Justice*. Oxford: Oxford University Press.

Riedl, A. and Van Winden, F. (2007). An experimental investigation of wage taxation and unemployment in closed and open economies. *European Economic Review*, 51(4):871–900.

Riedl, A. and van Winden, F. (2012). Input versus output taxation in an experimental international economy. *European Economic Review*, 56(2):216–232.

Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk: I. A definition. *Journal of Economic Theory*, 2(3):225–243.

Ruffle, B. J. (1998). More is better, but fair is fair: Tipping in dictator and ultimatum games. *Games and Economic Behavior*, 23(2):247–265.

Schlag, K. H. and van der Weele, J. (2013). Incentives for eliciting confidence intervals. *Available at SSRN: http://ssrn.com/abstract=2271061*.

Schnedler, W. and Vadovic, R. (2011). Legitimacy of control. *Journal of Economics & Management Strategy*, 20 (4):985–1009.

Schokkaert, E. and Capeau, B. (1991). Interindividual differences in opinions about distributive justice. *Kyklos*, 44(3):325–345.

Schokkaert, E. and Overlaet, B. (1989). Moral intuitions and economic models of distributive justice. *Social Choice and Welfare*, 6(1):19–31.

Sefton, M. (1992). Incentives in simple bargaining games. *Journal of Economic Psychology*, 13(2):263–276.

Selten, R. (1967). Die strategiemethode zur erforschung des eingeschränkt rationalen verhaltens im rahmen eines oligopolexperiments. *Beiträge zur experimentellen Wirtschaftsforschung, Heinz Sauermann (ed.), Vol. I, Tübingen: J.C.B. Mohr (Siebeck)*, 1:136–168.

Selten, R. and Ockenfels, A. (1998). An experimental solidarity game. *Journal of Economic Behavior & Organization*, 34:517–539.

Shayo, M. and Harel, A. (2012). Non-consequentialist voting. *Journal of Economic Behavior & Organization*, 81(1):299–313.

Siemens, F. A. V. (2013). Intention-based reciprocity and the hidden costs of control. *Journal of Economic Behavior & Organization*, 92(0):55 – 65.

Sliwka, D. (2007). Trust as a signal of a social norm and the hidden costs of incentive schemes. *American Economic Review*, 97(3):999–1012.

Tausch, F., Potters, J., and Riedl, A. (2013). An experimental investigation of risk sharing and adverse selection. *CESifo working paper no. 4192*.

The Washington Post (2012). Smokers, forces to pay more for health insurance, can get help with quitting. *Andrews M.*, January 2.

Thomson Reuters (2011). Paying for unhealthy behaviors. npr health poll. *October*.

Thral, R. and Rademacher, N. (2009). Bad luck vs. self-inflicted neediness - An experimental investigation of gift giving in a solidarity game. *Journal of Economic Psychology*, 30:517–526.

Tyran, J.-R. and Sausgruber, R. (2006). A little fairness may induce a lot of redistribution in democracy. *European Economic Review*, 50(2):469–485.

Van der Heijden, E. C., Nelissen, J. H., Potters, J. J., and Verbon, H. A. (1998). Transfers and the effect of monitoring in an overlapping-generations experiment. *European Economic Review*, 42(7):1363–1391.

Van der Lecq, S. and Steenbeek, O. W. (2007). *Costs and benefits of collective pension systems [electronic resource]*. Springer-Verlag Berlin.

Wakker, P. (2008). Explaining the characteristics of the power (crra) utility family. *Health Economics*, 17:1329–1344.

Wakker, P. (2010). *Prospect Theory for Risk and Ambiguity*. Cambridge: Cambridge University Press.

Yaari, M. E. and Bar-Hillel, M. (1984). On dividing justly. *Social Choice and Welfare*, 1(1):1–24.

Ziegelmeyer, A., Schmelz, K., and Ploner, M. (2012). Hidden costs of control: four repetitions and an extension. *Experimental Economics*, 15(2):323–340.

Acknowledgements

I am most grateful to my supervisors, Arno Riedl and Jan Potters. Working with them was not only instructive but also inspiring and fun. Thank you for your encouragement, support and valuable feedback!

I would further like to thank my co-author Elena for the productive and pleasant collaboration, as well as all my colleagues who unhesitatingly supported me in organizing and conducting experiments in the BEElab.

The presence of my friends inside and outside Maastricht has largely added to a wonderful past four years. Thank you all for your friendship that I appreciate greatly.

Last but not least I would like to thank my parents for their unconditional support in everything I do.

Franziska Tausch
Bonn, February 2014

Curriculum Vitae

Franziska Tausch was born in Wermelskirchen, Germany, on August 9, 1985. She obtained a diploma degree in Economics from the University of Bonn in 2010. During her studies she worked as a student research assistant in the department of energy regulation at the Scientific Institute for Communication Services (WIK) and the department of personnel and behavioral economics at the Institute for the Study of Labor (IZA).
At the end of 2009 Franziska started her dissertation research at Maastricht University. In 2013 she visited the Rady School of Management at the University of California San Diego and the Choice Lab at the Norwegian School of Economics. Her research has been presented at various international workshops and conferences.
Since February 2014 Franziska works at the Max Planck Institute for Research on Collective Goods in Bonn.

Franziska's research focuses on behavioral and experimental economics, decisions under uncertainty, risk sharing, coerced pro-social behavior and redistribution preferences.