PEF, **12** (3): 298–325, July 2013. © Cambridge University Press 2013
doi:10.1017/S1474747212000388 First published online 11 February 2013

Preferences for redistribution and pensions. What can we learn from experiments?

FRANZISKA TAUSCH

Department of Economics, Maastricht University, and Netspar
(*e-mail:* f.tausch@maastrichtuniversity.nl)

JAN POTTERS

Department of Economics, CentER, Tiber, Tilburg University, and Netspar

ARNO RIEDL

Department of Economics, Maastricht University, IZA, CESifo, and Netspar

Abstract

Redistribution is an inevitable feature of collective pension schemes and economic experiments have revealed that most people have a preference for redistribution that is not merely inspired by self-interest. However, little is known on how these preferences interact with preferences for different pension schemes. In this paper, we review the experimental evidence on preferences for redistribution and suggest some links to redistribution through pensions. For that purpose we distinguish between three types of situations. The first deals with distributional preferences behind a veil of ignorance. In the second type of situation, individuals make choices in front of the veil of ignorance and know their position. Finally, we discuss situations in which income is determined by interdependent rather than individual choices. In the closing sections of the paper, we discuss whether and how these experimental results speak to the redistribution issues of pensions. For example, do they argue for or against mandatory participation? Should we have less redistribution and more actuarial fairness? How does this depend on the type of redistribution involved?

Keywords: Experimental economics, redistribution preferences, pension systems, solidarity

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 centre 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 healthcare, unemployment insurance, disability insurance and poverty alleviation.

The remainder of the paper is organized as follows. Section 2 provides a brief introduction to the methodology of experimental economics. Section 3 is the main body of the paper, providing an extensive review of the experimental literature that deals with social preferences and redistribution. Section 4 gives a summary and our interpretation of the main results. Section 5 outlines the important missing elements in the experimental literature with respect to issues of pension solidarity. Finally, Section 6 concludes and presents the main implications.

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, 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 (Frechette, 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.

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 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 3.2. Finally, in Section 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.

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.

3.1.1. 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.

3.1.2. 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.

3.1.3. 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 aggregate, 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 result 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.

3.1.4. 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, 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 judgements 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 judgement 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 actively 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.

3.1.5. 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 by Cappelen *et al.* (2012) 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.

3.1.6. 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.

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.

3.2.1. 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 hypothetical. 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.

3.2.2. 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.

3.2.3. 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.

3.2.4. 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 programmes 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. Subjects 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).

3.2.5. 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 redistribution. 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.* (2011) 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.

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 outcomes 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?

3.3.1. 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.

3.3.2. 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 reconsider 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 voluntary 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.*, 1997). 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.

4. The main lessons

4.1 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.

4.2 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.

4.3 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.

4.4 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.

4.5 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.

4.6 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.

4.7 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.

4.8 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.

4.9 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.

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 analysing fairness ideas in such contexts was taken by Cappelen *et al*. (2012) 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.

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.

References

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.

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

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

Bolton, G. E. and Ockenfels, A. (2000) ERC: a theory of equity, reciprocity, and competition. *American Economic Review*, **90**: 166–193.

Cabrales, A., Nagel, R. and Rodriguez Mora, J. V. (2006) It is Hobbes, not Rousseau: an experiment on social insurance. UC3M Working Paper 07–08.

Camerer, C. (2003) *Behavioral Game Theory*. Princeton: Princeton University Press.

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

Cappelen, A. W., Konow, J., Sørensen, E. Ø. and Tungodden, B. (2012) Just luck: an experimental study of risk taking and fairness. Forthcoming in *American Economic Review*.

Carpenter, J., Verhoogen, E. and Burks, S. (2005) The effect of stakes in distribution experiments. *Economics Letters*, **86**: 393–398.

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.

Charness, G. and Genicot, G. (2009) Informal risk sharing in an infinite-horizon experiment. *Economic Journal*, **119**: 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, endowment uncertainty and risk sharing. Working Paper. Available online at http://users.monash.edu.au/~maitra/Risksharing_march26_2010.pdf.

Chavanne, D., McCabe, K. and Pia Paganelli, M. (2009) Redistributive Justice – Entitlements and Inequality in a Third-Party Dictator Game. George Mason University Working Paper.
Croson, R. and Gneezy, U. (2009) Gender differences in preferences. *Journal of Economic Literature*, **47**(2): 448–474.
Cox, J. C. and Deck, C. A. (2006) When are women more generous than men? *Economic Inquiry*, **44**(4): 587–598.
D'Excelle, B. and Riedl, A. (2010) Directed Generosity and Network Formation: Network Dimension Matters, IZA Discussion Paper No. 5356, December 2010.
Dickinson, D. L. and Tiefenthaler, J. (2002) What is fair? Experimental evidence. *Southern Economic Journal*, **69**(2): 414–428.
Durante, R. and Putterman, L. (2009) Preferences for Redistribution and Perception of Fairness: An Experimental Study. Mimeo. Brown University.
Eckel, C. and Grossman, P. (1998) Are women less selfish than men? Evidence from dictator experiments. *Economics Journal*, **108**: 726–735.
Esarey, J., Salmon, T. and Barrilleaux, C. (2009) *Social Insurance and Income Redistribution in a Laboratory Experiment*. Available online at http://ssrn.com/abstract = 1361134.
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**: 106–134.
Fehr, E. and Schmidt, K. (1999) A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, **114**: 817–868.
Forsythe, R., Horowitz, J., Savin, N. E. and Sefton, M. (1994) Fairness in simple bargaining experiments. *Games and Economic Behavior*, **6**(3): 347–369.
Frechette, G. (2011) *Laboratory Experiments: Professionals versus Students*. Available online at http://ssrn.com/abstract = 1939219.
Frohlich, N. and Oppenheimer, J. A. (1990) Choosing justice in experimental democracies with production. *American Political Science Review*, **84**(2): 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*, **31**(3): 606–637.
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. *Social Choice and Welfare*, **27**: 571–594.
Goeree, J., McConnell, M. A., Mitchell, T., Tromp, T. and Yarif, L. (2010) The 1/d law of giving. *American Economic Journal: Micro-Economics*, **2**(1): 183–203.
Güth, W., Offerman, T., Potters, J., Strobel, M. and Verbon, H. (2002) Are family transfers crowded out by public transfers? *Scandinavian Journal of Economics*, **104**(4): 587–604.
Hammond, P. (1975) Charity: altruism or cooperative egoism. In Phelps, E. (ed.), *Altruism, Morality and Economic Theory*. New York: Sage, pp. 115–131.
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. (2011) Inequality Aversion and Voting on Redistribution. Discussion Paper No. 11–18, Department of Economics, University of Copenhagen.
Hoffman, E., McCabe, K. A., Shachat, K. and Vernon, L. S. (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. *American Economic Review*, **86**(3): 653–660.
Jakiela, P. (2009) How fair shares compare: experimental evidence from two cultures. Working Paper.

Johannesson, M. and Persson, B. (2000) Non-reciprocal altruism in dictator games. *Economics Letters*, **69**: 137–142.
Klor, E. and Shayo, M. (2010) Social identity and preferences over redistribution. *Journal of Public Economics*, **94**: 269–278.
Konow, J. (1996) A positive theory of economic fairness. *Journal of Economic Behavior and Organization*, **31**: 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. (1988) Social contracts as assets: a possible solution to the time-inconsistency problem. *American Economic Review*, **78**: 662–677.
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.-A. (2009) Directed altruism and enforced reciprocity in social networks. *Quarterly Journal of Economics*, **124**(4): 1815–1851.
List, J. A. (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 and 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**: 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**: 308–317.
Normann, H.-T. and Ricciuti, R. (2009) Laboratory experiments for economic policy making. *Journal of Economic Surveys*, **23**: 407–463.
Offerman, T., Potters, J. and Verbon, H. (2001) Cooperation in an overlapping generations experiment. *Games and Economic Behavior*, **36**: 264–275.
Oxoby, R. and Spraggon, J. (2008) Mine and yours: property rights in dictator games. *Journal of Economic Behavior and Organization*, **65**: 703–713.
Paetzel, F., Sausgruber, R. and Traub, S. (2012) *Social Preferences and Voting on Reform – An Experimental Study*. Available online at http://ssrn.com/abstract=2066455.
Rawls, J. (1971) *A Theory of Justice*. Oxford: Oxford University Press.
Riedl, A. and van Winden, F. A. A. M. (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. A. A. M. (2012) Input versus output taxation in an experimental international economy. *European Economic Review*, **56**(2): 216–232.
Ruffle, B. J. (1998) More is better, but fair is fair: tipping in dictator and ultimatum games. *Games and Economic Behavior*, **23**: 247–265.
Schokkaert, E. and Capeau, B. (1991) Interindividual differences in opinions about distributive justice. *Kyklos*, **44**: 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**: 263–276.
Selten, R. and Ockenfels, A. (1998) An experimental solidarity game. *Journal of Economic Behavior and Organization*, **34**: 517–539.
Shayo, M. and Harel, A. (2012) Non-consequentialist voting. *Journal of Economic Behavior and Organization*, **81**: 299–313.
Tyran, J.-R. and Sausgruber, R. (2006) A little fairness may induce a lot of redistribution in democracy. *European Economic Review*, **50**(2): 469–85.

Yaari, M. E. and Bar-Hillel, M. (1984) On dividing justly. *Social Choice and Welfare*, **1**: 1–24.
van der Heijden, E., Nelissen, J., Potters, J. and Verbon, H. (1997) Intergenerational transfers and private savings: an experimental study. *Kyklos*, **50**(2): 207–220.
van der Heijden, E., Nelissen, J., Potters, J. and Verbon, H. (1998) Transfers and the effect of monitoring in an overlapping-generations experiment. *European Economic Review*, **42**: 1363–1391.
van der Lecq, S. and Steenbeek, O. (2007) Introduction. In van der Lecq, S. and Steenbeek, O. (eds), *Costs and Benefits of Collective Pension Systems*. Berlin: Springer-Verlag.