Stockholm Institute of Transition Economics

WORKING PAPER

August 2009

No. 3

Purely Procedural Preferences – Beyond Procedural Equity and Reciprocity

Nadine Chlaß, Werner Güth, and Topi Miettinen

STOCKHOLM INSTITUTE OF TRANSITION ECONOMICS

Stockholm Institute of Transition Economics (SITE) · Stockholm School of Economics · Box 6501 · SE-113 83 Stockholm · Sweden

Purely Procedural Preferences
- Beyond Procedural Equity and Reciprocity -

Nadine Chlaß* Werner Güth** Topi Miettinen‡

March 2, 2016

Abstract

We show theoretically that all existing procedural preference models to date are ultimately based on outcomes. We design pairs of procedures such that – according to all of these preference models – parties should be indifferent between the procedures at hand. In particular, the procedures should yield the same outcomes, the same expected outcomes and carry the same information on parties' intentions. We conduct an experiment and find that individuals do, however, still reveal preferences. To explore why this happens, we elicit individuals' ways to make moral judgements in a standardized psychological test. The preferences we find link to the degree to which individuals invoke the equality of basic rights and liberties in their moral judgement – an ethical ideal not yet captured by any preference model. We illustrate that individuals who hold such ideals compensate the perceived unfairness of a procedure behaviourally and offer formalizations for such preferences.

JEL Classification: C90, C78, D02, D63

Keywords: procedural preferences, experiment, institutional design

*Departments of Economics, University of Turku, FI-20014 Turku, Finland and the University of Jena, Carl-Zeiss-Str. 3, 07743 Jena, Germany. Email: nadine.chlass@utu.fi. Phone: 00358 (0)2333 9112.

**Department of Economics and Finance, LUISS Guido Carli, Viale Pola 12, 00198 Rome, Italy, and the Max Planck Institute for Research On Collective Goods, Kurt-Schumacher-Straße 10, 53113 Bonn, Germany.

‡Hanken School of Economics, P.O. Box 479, Helsinki Fi-00101, Finland. We thank participants of the EEA/ESEM, the ESA Europe meetings, and participants of seminars at the Max Planck Institute of Economics in Jena, and the University of East Anglia for discussions. We thank Christoph Göring for programming the z-Tree code. Topi Miettinen gratefully acknowledges financial support from the Yrjö Jahnsson foundation.

1 Introduction

In some areas of life, procedures may be vitally important when they do not have even a stochastic influence on outcomes. In an election, for instance, great care is taken to grant each individual an equal opportunity to vote, to make the voting simple, and to elect a candidate in a transparent way. Yet, one may plainly refuse to acknowledge a candidate's victory, if it is learned that the election violated some of the criteria mentioned before. Notably, such a concern may be independent of any potential outcome.

Since Thibault and Walker's (1975) seminal contribution, an impressive body of research in psychology – and more recently, also in economics – has studied the topic of procedural fairness. Procedural justice is best understood if contrasted with *distributive justice*. While distributive justice (Adams 1965) is concerned with unjust allocations and human reactions to these, *procedural justice* explores the fairness of the principles and measures taken to reach such allocations and how individuals react to the application of these principles.

Procedural fairness is a necessary building block for economic prosperity and a stable society. A third party resolving a property rights dispute, for instance, needs legitimacy for its authority. This legitimacy springs ultimately from a shared perception between the dispute parties and outsiders about the fairness of the procedures employed (Lind 2001; Tyler 2004). Perceived process fairness also promotes compliance by the dispute parties to the verdicts of the authority (Lind 2001). Since the seminal work of Thibaut and Walker (1975), research in psychology (Lind 2001; Tyler 2004) and experimental and behavioural economics (Falk et al. 2003; Bolton et al. 2005; Brandts et al. 2006) have come to establish and support these views.

Psychological and economic research into procedural fairness employ different methods to strive for overlapping but different goals. Both disciplines have sought to disentangle process fairness from distributive fairness. Cropanzano and Ambrose (2001, pp. 125) state that in psychology, distributive justice is operationalized as "individuals' reactions to economic or quasi-economic allocations", and procedural justice as "individuals' reactions to the allocation of socioemotional benefits".

Economists resort to game theoretic models to clarify the difference between distributive (outcome-related) and procedural fairness. So far, the overarching principle in the procedural fairness literature has been to compare the outcomes of a procedure to the outcomes other procedures would have generated. When the fairness value of an outcome depends on counterfactual paths of the process (or the game), then the process itself matters, not only the outcome. By exploiting this principle, a first strand of research studies whether and how individuals discriminate between fair and unfair explicit randomizations over outcomes (Karni and Safra 2002; Bolton et al. 2005; Karni et al. 2008; Trautmann 2009; Krawczyk 2011; Krawczyk and LeLec 2010; Kircher et al. 2009)[1]. A second strand

[1]Güth and Tietz (1986) or Grimalda (2012) study equal and unequal randomizations over the roles of an ultimatum game. Formally, this translates into comparing fair and unfair lotteries over unequal outcomes (the proposer usually has a higher payoff than the responderin the ultimatum game). Proposers make higher offers when the lottery is unfair than when it is fair; responders reject higher offers when the lottery is unfair than when it is fair. These behavioural compensations are analogous to those found in (Bolton et al. 2005). Mertins and Albert (2015) find that responders who can directly affect proposers' take rate in a power-to-take game destroy the pie less often than responders who cannot

of research explores how kind (economically generous) an individual deems an opponent's choice of a specific process[2] compared to what she believes would have happened had the opponent chosen a different process (Rabin 1993; Blount 1995; Dufwenberg and Kirchsteiger 2004; Falk and Fischbacher 2006; Brandts et al. 2006; Sebald 2010; Aldashev et al. 2010)[3]. Both strands have made important contributions to understanding differences in individual reactions to procedures and to advancing the optimal design of institutions and organizations. Yet, from a bird's-eye perspective, it is evident that in both strands, individuals evaluate the fairness of a given process by their subjective expectations of the social and economic benefits which alternative processes would have generated. Therefore, both strands of research ultimately refer to the distribution of outcomes and thus stress that the rules of the game matter *because* they impact the distribution of outcomes.

Similar connections between distributional and process fairness have also been admitted in psychology. Two leading figures in the field – Cropanzano and Ambrose – summarize the decades of psychological research into procedural and organizational justice (2001, pp. 119-120) concluding that:

"the procedural justice and distributive justice are more similar than is generally believed... both procedural justice perceptions and distributive justice perceptions are, in some sense, derived from individuals' expectations about outcomes."

Thus, whether looking into economics or psychology, procedural fairness tends to be evaluated applying yardsticks for distributive fairness.

In this paper, we study process fairness which does not refer to the outcomes of procedures, or to the outcomes of alternative procedures at all – process fairness which requires that the rules of the game do not privilege any party in terms of available information or decision rights. Individuals may disapprove of such privileges even if these do not assist the privileged party in the pursuit of her material ends. To test this idea, we let subjects choose between two pie-splitting procedures. These procedures are designed such that all social preference models predict subjects have the same material and social payoffs in each of the two pie-splitting procedures. We do not only elicit individual choices *between* procedures but also actual beliefs and behaviour *within* each procedure to verify that subjects are indeed indifferent between the procedures – even out of equilibrium. Even when actual beliefs and actions indicate indifference, we find that subjects are willing to pay for having the allocation determined by one procedure and not the other. This is our first main finding.[4]

We also study the rationale behind *the observed choices between procedures* and find that they are morally motivated. Relying upon Jean Piaget's (1948) and Lawrence Kohlberg's (1969, 1984) work,

affect the take rate. Formally, the power-to-take games with and without responder influence on the take-rate translate into lotteries with different expected behaviour – and hence, different expected outcomes.

[2]By 'process' we mean a 'path' in the extensive form of a game.

[3]The experimental findings of (Brandts et al. 2006) and (Albrecht and Mertins 2015) can be understood from this perspective. In (Sebald 2010) and Aldashev et al. (2010) players may explicitly randomize at her own decision nodes of the game when choosing between actions, and opponents can hold beliefs that actions were so chosen.

[4]The type of procedural preference we study corresponds to Class I of procedural utility proposed by Benz and Stutzer (2003): preference from institutions per se. Yet, we do not elicit what kind of institutions/procedures subjects view as ideally fair. We merely elicit pairwise preference comparisons between two alternative procedures.

we elicit in a standardized moral judgement test (Georg Lind 1978, 2000, 2008) how an individual arrives at the conclusion that some portrayed action is either right or wrong. More specifically, we elicit by which degree individuals refer to expectations about punishment or reward, social norms, others' expectations or intentions when judging whether some course of action is right or wrong. We obtain a measure for individuals' preferences over the set of moral ideals which are at the heart of all current social preference models in economics. Kohlberg's field work lists yet two other moral ideals whose consequences have not yet been modelled in economics: the ideal that each member of society enjoy the same rights – justified either in terms of the social contract, e.g. (Rawls 1971; Binmore 1994) or the existence of some inalienable human right. We find that the extent to which subjects invoke these two last moral ideals in their moral judgement increases the likelihood that subjects indicate a preference between our two pie-splitting games. This is our second main finding. The second result supports the conclusion that the procedural choices we observe do not reflect mistakes in decision making, or outcome-related differences between the procedures which we cannot measure.[5]

Already the early literature on democracy and its federal organization points out that procedures which grant all parties equal decision rights and information can become undesirably complex and hard to handle (de Tocqueville 1868, p. 276). Hence, the decision maker may need to trade off potentially conflicting properties when choosing between alternative procedures. To illustrate such trade-offs, take the following example. Individuals may prefer to grant every agent an option to (dis)agree on some proposal rather than denying one or several agents their say. Individuals may also prefer that every agent is properly informed about that proposal before she opts to agree or to disagree. Incorporating these aspects into a procedure will increase agents' participation and freedom of choice. At the same time, these aspects increase the potential number of instances where an agreement cannot be found – and the need for an expedient regulation of those. In the real world, choices of institutions tend to involve such trade-offs. Therefore, we let subjects decide between alternatives which are predicted to induce identical expectations and outcomes. In the appendix, we offer some examples of privileges which a procedure can grant, and how to formally capture the corresponding degree of purely procedural (un)fairness without referring to outcomes in any way.[6] Purely procedural preferences are economically and politically relevant. Survey studies suggest that fair procedures catalyze the smooth functioning of organizations and institutions. The more an institution or an organization is deemed to employ fair procedures, the higher are organizational engagement, performance, and cooperation

[5]We opt for a simple yet conservative belief elicitation method (Schlag and Tremewan 2012, Schlag et al. 2014), conservative in that it biases beliefs which express certainty (indifference) downward to uncertain beliefs. We countercheck the accuracy of the indifference condition by controlling that also a new normative ideal underlies subjects' choices.

[6]We discuss four purely procedural concerns which may capture the observed choices over procedures: (i) a concern for equal opportunities (freedom of choice), (ii) a concern to rule out unkind opportunities, or a concern for efficiency, (iii) a concern for symmetric information, and (iv) a concern for procedural simplicity. Procedural fairness research in organizational psychology enlists six properties that fair procedures should conform to: (i) consistency (with equal opportunity as an integral subproperty), (ii) freedom from bias, (iii) accuracy (all relevant information is available when decisions are taken), (iv) correctability, (v) representativeness (of the parties' interests , often coined as "voice"), (vi) compliance with prevailing ethical standards (Leventhal 1976; Cropanzano and Ambrose 2001). Sociologist Max Weber uses the first three aspects to define how much *power* a party holds. In (Weber 1921 I §16), power is about the number of opportunities to implement one's will, also against opposition. Moreover, it arises from the fact that information is kept asymmetrically to a small circle of people close to the decision maker (Weber 1921, X §3).

(Tyler 2000; Sondak and Tyler 2007; De Cremer et al. 2005). Frey and Stutzer (2005) find that inhabitants of Swiss Cantons with greater democratic participatory rights are more satisfied with their lives. Thereby, life satisfaction does not only increase because participation is seen to improve the outcomes of political decision making – self-reported life satisfaction also increases in the right of proper participation itself. These surveys suggest that better insight into procedural fairness might also benefit institutional design.

Economists have indeed called for more economic and behavioural research into the relatively new field of procedural fairness (Rabin 1993; Rabin 2002; Konow 2003; Engelmann and Strobel 2004). Even experimental economics which has long taken intrinsic fairness concerns seriously, has only recently turned to analyze the effects of procedural aspects. To date, however, experiments on procedural preferences invariably compare procedures which induce outcome distributions with different statistical moment(s) – the preferences studied can thus be about those differences. Exceptions are Fehr et al. (2013) or Bartling et al. (2014) who show in studies on control rights and delegation that principals have intrinsic value for maintaining the power to decide and control and that reducing control or delegating decision rights to an agent has a positive welfare effect beyond what reciprocation can explain (Charness et al. 2012). We, instead, report evidence for *ethical* concerns about the *distribution* of rights across players, and a concern for procedural efficiency – parties remove their veto if they expect that it cannot change payoffs and if they can still voice their opinion without veto.

The following section describes the two-player pie-division procedures we use. Section 3 verifies that all relevant preference models and theories predict procedurally invariant outcomes within each pair of procedures – in, but also out of equilibrium. Section 4 introduces our experimental design and the experimental test used to describe individuals' moral judgement. Section 5 presents our main results, Section 6 cross-checks some aspects of cleanliness of our design. Section 7 concludes and argues that the purely procedural preferences we report may resolve controversies about other preference types.

2 Allocation procedures

We design three simple procedures which generate the same outcomes, the same expected payoffs, and the same psychological payoffs according to a variety of social preference theories. This section describes the procedures we use, the next section discusses their outcome-invariance. Let 200 units be shared among two parties. One party, the proposer (P), has more allocation power than the other, the responder (R). Two divisions of the pie are possible; a fair one, where both the proposer and the responder obtain 100 units and an unfair one where the proposer obtains 20 units and the responder 180 units. Thus, the unfair allocation favors the less powerful responder. We introduce three procedures for sharing the 200 units in either way: a mini dictator game, a mini yes-no game (Gehrig et al. 2007), and a mini ultimatum game (Güth et al. 1982).

The first procedure, a dictator game (DG), leaves the responder R no option to choose in a payoff-relevant way. Whatever proposer P chooses is implemented. In our specific setting, the responder

can agree or disagree with the proposal but her choice does not affect the outcome[7]. The DG is thus a one person decision problem in a two-person environment. A second procedure, the yes-no game (YNG), grants the responder an *unconditional* opportunity to choose. P proposes either (100,100) or (20,180) and R decides whether to accept without knowing the proposal made by P. Hence, R cannot condition her decision on P's proposal. If R agrees, the proposal is implemented. If she rejects, both parties earn zero payoffs. Therefore, the yes-no game is a two-player game with each player having two options only.

A third procedure, the ultimatum game (UG), grants the responder a *conditional* opportunity to act. As in the yes-no game, P proposes one of the two allocations. R decides for each potential proposal whether to accept or to reject it. Again, a rejection leads to zero payoffs whereas acceptance implements the proposed sharing.

We confront each subject with one pair of alternative procedures to choose from. Each subject chooses either between the yes-no game and the Ultimatum game, or between the Ultimatum game and the Dictator game. The details of the design are explained in Section 5.

3 Predictions within procedures

In this section we show that the games and the monetary payoffs in section 2 were chosen such that central preference theories predict the same equilibrium outcome, equilibrium behaviour, and the same equilibrium beliefs[8] in each allocation procedure. Table 1 summarizes the results. Experimentally, however, individuals may yet not always comply with (a) the predicted equilibrium behaviour, and (b) the predicted equilibrium beliefs. For these cases, we show that subjects are still indifferent *if* they show identical behaviour in each allocation procedure, and expect the same opponent action *with certainty* in each procedure[9]. Put differently, subjects are indifferent if they do not expect to learn anything about their opponent in any procedure when choosing the allocation procedure. Table 2 reviews which conditions ensure that on and off equilibrium, the procedures generate distributions of outcomes with identical statistical moments under a given preference model.

To control therefore whether participants in our experiment are indeed indifferent between two procedures, we elicit *actual actions* and *actual beliefs* in each procedure. In section 6.1, we will impose the most restrictive indifference condition from table 2 on subjects' actions and beliefs before extending our analysis to the entire set of subjects. In section 6.2 we further show that the *motivation* behind

[7]We introduce this feature to reduce the number of differences across procedures. Since all three procedures give responders a 'voice', she can always express her opinion, and her preference for a specific procedure cannot be motivated by a desire to express her opinion. In our setup, the responder's voice is not communicated to the proposer, though.

[8]Throughout, the solution concepts applied in table 1 postulate sequential rationality.

[9]To see why this condition is necessary, take a player with reciprocal preferences. She would – when experiencing kindness (unkindness) – wish to be kind (unkind) in return. In the ultimatum game, a responder may learn about the proposer's kindness during play, and thus better reciprocate the proposer's actions in order to yield higher psychological payoffs [Reciprocal models – unlike standard game theory – assume that payoffs are quadratic, and not linear in beliefs. This is necessary to allow for an emotional payoff term which is the product of players' kindness towards each other. Thus, the choice probabilities will be squared in the payoff expressions]. This may be a reason to prefer the ultimatum over the yes-no game where parties cannot learn about others' kindness.

subjects' procedural choices under this condition is at odds with all preference models discussed here.

Let us now illustrate that existing and ultimately outcome-based preference models have a hard time explaining procedural preferences in this paper's setting.

3.1 Distributive theories

Self-interested opportunism. If R is opportunistic, she only cares about her share of the 200 units of pie and never rejects any proposal. Anticipating R's opportunism, P selects the allocation (100,100) in all three games and R accepts whenever she has the opportunity.[10] The expected payoff in each procedure is 100 for each player. Self-interested players are therefore indifferent between all three allocation procedures. *Self-interested parties who violate these predictions are still procedurally indifferent if their actual behaviour, and actual beliefs are the same in all procedures.*

			BEHAVIOURAL PREDICTIONS						
			make fair proposal DG	make fair proposal YNG	make fair proposal UG	accept in YNG	accept (100,100) UG	accept (20,180) UG	**same outcomes across procedures**
SOCIAL PREFERENCE MODELS	Outcome based	Self Interest	+	+	+	+	+	+ off eq.path	+
		Inequity Aversion	+	+	+	+	+	+ off eq.path	+
		Altruism	depends on degree of altruism	depends on degree of altruism	depends on degree of altruism	+	+	+	+
	Reciprocity based (PGT)	Falk & Fischbacher (2006)	+	+	+	+	+	+ off eq.path	+
		Dufwenberg& Kirchsteiger (2004)[11]	+	+	+	+	+	+ off eq.path	+
	Guilt based	Battigalli & Dufwenberg (2007)	depends on sensitivity to guilt	depends on sensitivity to guilt	depends on sensitivity to guilt	+	+	+	+
PROCEDURAL FAIRNESS MODELS	Inequity based	e.g. Bolton et al. (2005)	+	+	+	+	+	+	+
	Reciprocity based	Sebald (2010)	+	+	+	+	+	+	+

Table 1: **Predictions of central social preference theories.**

[10] These strategies are sequentially rational (Selten 1967).

[11] See appendix E.

Inequity aversion. Models of allocative fairness (Bolton 1991; Bolton and Ockenfels 2000; Fehr and Schmidt 1999) assume that a player's utility does not only increase in a player's private payoff, but also in the equality of payoffs. Fehr and Schmidt (1999) assume that each player's own payoff and her payoff from (in)equality are additively separable. That is, if a player earns x units and her opponent earns y units, then the player's utility is $x - a \times max\{(y-x), 0\} - b \times max\{(x-y), 0\}$ where a and b denote non-negative individual parameters. Further, the model assumes that players suffer more from disadvantageous than from advantageous inequality, that is, $a \geq b$. A player strictly prefers the allocation $(0,0)$ to (x,y) with favourable inequality $x > y$ iff $b > \frac{x}{(x-y)}$. A player strictly prefers $(0,0)$ to the allocation (x,y) with unfavourable inequality $x < y$ iff $a > \frac{x}{(y-x)}$. For our two allocations ($x = 100$, $y = 100$) and ($x = 180$, $y = 20$), inequity averse responder with $b < 1$ would accept all proposals. If so, inequity-averse proposers maximize their utility by proposing (100,100). The expected payoff is 100 for each player in each procedure. Thus, neither player should prefer one procedure over another. *Inequity-averse parties who for some reason, violate these predictions are still procedurally indifferent if their actual behaviour, and their actual beliefs are the same in all procedures.* An inequity averse individual invokes a social reference point about the distribution of material payoffs (Fehr and Schmidt, pp. 820-821, Bolton and Ockenfels, p. 172), or put differently, a social norm about the equality of outcomes (Bolton et al. 2005, p. 1068) to derive the right course of action.

3.2 Psychological game theory

Reciprocity. If responders care for the kindness of the intention behind a proposal, they compare the actual proposal with other proposals that could have been made. The kindness of a proposal therefore depends on the set of possible proposals. The unrestricted set of proposals is a set where the pie can be split into any numerically possible way. On this set, the equal division is fair. If only two options are available, the equal split may be considered even fairer. Indeed, Falk et al (2003) hardly ever find responders who reject meager offers in mini-ultimatum games when only two proposals are possible – suggesting that even meager offers are more acceptable for the smaller set. Apart from restricting the set of proposals, our experimental design also has no proposal where the proposer earns more than the responder. Hence, both allocations: $(100, 100)$, and $(20, 180)$ should appear kind and be accepted. We next discuss reciprocal concerns in the frameworks of Falk and Fischbacher (2006), and Dufwenberg and Kirchsteiger (2004).[12] Throughout, reciprocal preference models assume that individuals invoke others' intentions to derive the right course of action.

Reciprocity – Falk and Fischbacher (2006). The kindness of player j towards i at node n is defined as $\varphi_j(n, s_i'', s_i') := \vartheta_j(n, s_i'', s_i')\Delta_j(n, s_i'', s_i')$ where s_i' represents i's first-order belief about the strategy of j and s_i'' is i's second-order belief (the belief about the first-order belief of j). In equilibrium, this second-order belief coincides with a player's actual behaviour. The term $\Delta_j(n, s_i'', s_i') = x_i(n, s_i'', s_i') -$

[12]Cox et al. (2007, 2008) formulate an alternative to the psychological game theory models of reciprocity discussed in the main text. In their model, a player's lost or gained payoff opportunities at earlier nodes of an extensive form game influence the subsequent marginal rate of substitution (MRS) between the player's own earnings and those of her opponent. The MRS remains constant across two games where the fair proposal is always proposed and each proposal is always accepted. Thus, also according to Cox et al. (2007, 2008) players are indifferent between this paper's protocols.

$y_j(n, s_i'', s_i')$ expresses the perceived payoff difference, $\vartheta_j(n, s_i'', s_i') \in [0, 1]$ measures the degree of intentionality in j 's choices. For negative Δ_j, player j is unkind to i whereas for positive Δ_j, player j is kind. For binary choices, a player is intentionally unkind if she gives her opponent a smaller share of the pie than she keeps herself when she might have offered the opponent the larger share. A player is unintentionally unkind to her opponent if she gives her opponent a smaller share of the pie than she keeps for herself but had no opportunity to give the same or the larger share. For all our procedures and all their outcomes, the difference between what the proposer gave and what she kept, i.e. Δ_j, remains non-negative. Therefore, the *proposer* cannot be unkind.

The *responder* ensures equal payoffs both if she accepts the fair offer, and if she rejects it. The fair proposal $(100, 100)$ is not unkind and is therefore always accepted. The generous proposal $(20, 180)$ is even kinder. If a responder accepts this generous offer, she is unkind – because this gives her opponent less than herself. However, this unkindness is not deemed intentional, since rejecting the generous offer would give the proposer even less than the generous proposal does. Thus, the generous offer is accepted provided that purely distributional motives do not matter. If, however, an individual holds a high concern for equal outcomes *and* sufficiently strong reciprocal motives, Falk and Fischbacher (2006) can predict rejections of the generous offer in equilibrium. This reaction to the generous offer does, however, not matter, since the proposer in equilibrium prefers to propose the fair offer anyway. The fair proposal is accepted with certainty in every perfect equilibrium of both the mini ultimatum and the mini yes-no game. In the dictator game, the responder cannot be intentionally kind or unkind since she has no influence on any payoff. The proposer thus chooses the fair proposal. In summary, Falk and Fischbacher (2006) predict that the fair offer is always proposed and accepted with certainty in all procedures, and that each player earns 100. Since there are no payoff differences, the psychological payoffs are zero and the equilibrium payoffs identical in all procedures. No player should prefer one procedure over another.

Reciprocity – Dufwenberg and Kirchsteiger (2004). This model of reciprocity first identifies efficient strategies. The difference between the payoff a player gives her opponent with a specific strategy and the average payoff a player gives her opponent over all efficient strategies which are still available at a given node measures the kindness of a specific strategy (see Dufwenberg and Kirchsteiger, pp. 276). In every protocol of our setting, there is a single efficient responder strategy: the pure strategy which accepts every proposal. Thus, all responder strategies that put a positive probability on rejection are unkind, and the responder can only be neutral or unkind towards the proposer. This implies that the proposer always prefers the fair offer if the probabilities of acceptance of each offer are equal: there is no kindness she would need to reciprocate. Knowing that the fair offer will be proposed for sure, the kindness of the responder who rejects with probability q equals $q \cdot 100$ for the yes-no game, and the ultimatum game. If the proposer believes that each offer is accepted with probability q, her kindness in proposing the fair offer is[13] $(q \cdot 100 - q \cdot (100 + 180)/2)$ in both games. Each player's equilibrium payoff is thus identical in the mini-ultimatum and the mini- yes-no game given her sensitivity to reciprocity.

[13]The difference between the expected responder payoff in the fair offer, i.e. $q \cdot 100$, and the expected average responder payoff over all efficient available strategies, i.e. $q \cdot (100 + 180)/2$.

In equilibrium therefore, players are indifferent between these two procedures.

In the dictator game, each proposal is accepted with certainty. The responder has no influence on payoffs and for this reason, is always neutral towards the proposer. Therefore, psychological payoffs are zero, preferences coincide with rational self-interest, and the proposer chooses the fair proposal. As we saw above for the ultimatum and yes-no game, accepting both offers with certainty is efficient and expresses zero kindness towards the proposer. The psychological payoffs are zero as in the dictator game. Players who believe that every proposal is accepted with certainty in all games and who expect the fair proposal to be always proposed are indifferent between the dictator, ultimatum, and yes-no game. In appendix E, we characterize all equilibria of the games at hand under the constraint of equal acceptance probabilities across nodes and games (which is a necessary condition for procedural indifference and a feature imposed by the empirical analysis).

General remark on psychological games. In psychological games, payoffs depend explicitly on beliefs and thus, expected payoffs do not have to be linear in probabilities (contrary to standard expected utility theory). Specifically, the psychological payoffs of the two theories of reciprocity are quadratic in beliefs. For instance, the responder's evaluation of the proposer's kindness depends explicitly and quadratically on how likely she deems the generous offer. We denote this probability by $1-p$. Since in the ultimatum game, the responder reacts to updated information about this probability, the expected payoff of the responder differs from his expected payoff in the yes-no game (where the responder does not receive an information update) whenever the ex-ante belief about the probability of the fair offer is $0 < p < 1$, even if ex ante beliefs are identical in the two games (by Jensen's inequality). The expected payoffs are yet equal in the two games if ex ante, the fair offer is either certain, i.e. $p = 1$, (as predicted by sequential reciprocity equilibrium if acceptance rates are equal, see appendix E) or impossible, i.e. $p = 0$.

Guilt aversion (Battigalli and Dufwenberg 2007; Charness and Dufwenberg 2006) is yet another other-regarding concern which can also be modelled via psychological game theory. In these theories, guilt matters only if a player harms the other and lets the other down (Bicchieri, 2006, pp. 52; Battigalli and Dufwenberg, 2007, pp. 171; Miettinen, 2013, pp. 71). If the responder expects the proposer to expect rejection, the responder does not harm the proposer by accepting instead and the responder's guilt payoff is zero. Thus, the responder's preferences coincide with rational self-interest and she always accepts. If the responder expected the proposer to put some weight on acceptance in her beliefs, rejecting would harm the proposer. The responder's guilt payoff will then only increase her incentive to accept. Therefore, the responder always accepts, and her guilt payoff is zero. A very guilt averse proposer who very much expects the responder to expect a generous offer might indeed offer $(20, 180)$. However, as long as actual actions and actual beliefs are the same for two procedures, guilt averse parties are indifferent between them. This differs from reciprocity, because in guilt aversion, psychological payoffs are linear in beliefs (Battigalli and Dufwenberg 2007), and not quadratic. In terms of normative ideals, a guilt averse individual invokes others' expectations (Battigalli and Dufwenberg 2007, p. 170) or social norms (Bicchieri 2006, López-Pérez 2008) to derive the right course of action.

3.3 Economic models of procedural fairness

Recently, *economic approaches to procedural fairness* have been developed, some building upon inequity aversion (Bolton et al. 2005; Krawczyk 2011; Trautmann 2009), others upon reciprocity (Sebald 2010)[14]. Even these approaches predict indifference between the two pie-sharing games in each of the two pairs of games. Bolton and Ockenfels (2005) formulate that individuals are inequity-averse over expected payoffs and prefer lotteries with similar expected payoffs for both players to lotteries with dissimilar expected payoffs. Applying this – or the other two inequity based models of procedural preferences (Trautmann 2009; Krawczyk 2011) – to our setting, we find that participants who hold the same beliefs in two procedures will also expect the same payoffs in each procedure and therefore, be indifferent between the procedures.

Sebald (2010) allows the preference to be influenced by the kindness of a procedure, that is, the kindness the opponent would have shown had she chosen that procedure. In Sebald's model – contrary to Dufwenberg and Kirchsteiger (2004) – the responder does not update her beliefs about the proposer's choice probabilities in the ultimatum game when she learns the proposal that has been made (if both proposals have a positive probability ex ante). Thus, if a player has procedurally invariant actions and beliefs, she is predicted to be indifferent between the mini yes-no game and the mini ultimatum game. Similarly, if each proposal is accepted for sure in the ultimatum game, the responder is neither kind nor unkind towards the proposer (recall that accepting is the only efficient strategy) and the psychological payoffs are always zero in the dictator, and the ultimatum game. Thus, if each proposal is proposed with equal probability in these games, players are indifferent. Table 2 reviews the conditions under which participants are procedurally indifferent.

3.4 Summary

In summary, economic approaches to procedural fairness are – just as their non-mathematical counterparts in psychology (Cropanzano and Ambrose 2001) – based on distributive fairness. They predict players to be indifferent between our three allocation procedures since these procedures induce identical distributions of outcomes. Thus, if we still observe preferences for one allocation procedure over another, this would suggest a new type of procedural preference. In order to observe such novel preferences, (i) the procedures must vary in aspects which are meaningful to the subject, and (ii) behaviour and beliefs must be invariant in each of the two procedures. This is why (i) we constrain our interest to the mini-games with only a fair and a generous offer, (ii) why we elicit behaviour and beliefs in each procedure, and (iii) why we let each player make choices in both roles such that participants exert maximal cognitive effort to put themselves into the shoes of the other player to understand that fair proposals are likely, and that both offers are acceptable.

If the choices between our procedures do indeed spring from a novel preference, a moral ideal should be at play which does not underlie any of the preference types we have discussed. To check this, we will elicit individuals' moral preferences – the moral criteria which they employ to judge about the

[14]Sebald's model is based upon the reciprocity model of Dufwenberg and Kirchsteiger (2004).

right or wrong of an action – and try to explain the procedural choices we observe in sections 6.2 and 7. These choices should *neither* link to individuals' propensity to invoke others' expectations, nor to others' intentions, nor to social norms or references, nor to status. Indeed, choices should not link to any of the ethical ideals underlying the preferences discussed here in section 3. Next, we briefly illustrate this paper's ideas how one might formulate purely procedural preferences which do not rely on distributive fairness before we describe the experimental design.

				PLAYERS ARE PROCEDURALLY INDIFFERENT...	
				...in perfect equilibrium.	...off equilibrium... ...if they choose analogous pure strategies AND deem some outcomes certain/impossible in all games.
SOCIAL PREFERENCE MODELS	OUTCOMES		Self interest	+	+
			Inequity Aversion	+	+
			Altruism	+	+
	INTENTIONS	Reciprocity	Falk & Fischbacher (2006)	+	+
			Dufwenberg & Kirchsteiger (2004)	+	+
		Other	Guilt	+	+
PROCEDURAL PREFERENCE MODELS		Inequity-based	Bolton et al. (2005) Trautmann (2009) Krawczyk (2011)	+	+
		Reciprocity-based	Sebald (2010)	+	+

Table 2: **When are players procedurally indifferent in, and off equilibrium?**

4 Purely Procedural Preferences

Suppose now that a player believes that all games do indeed lead with certainty to the same allocation and that hence, all preference models above predict indifference – both in, and off equilibrium. Suppose that moreover, none of the normative ideals upon which existing preference models build are at play. Which type of preferences could a player still hold? To date, behavioural economics knows various intriguing formalizations of the idea that people care how material benefits are distributed across

individuals (Fehr and Schmidt 1999; Bolton and Ockenfels 2000). In this paper, we pursue the idea that people might care about such things as *how rights are distributed across parties*. One could for instance, express a player's decision rights in a procedure by the cardinality of her strategy set – or analogously, the cardinality of her action set summed over all instances where she is called upon to play[15] – counting only strategies which generate genuinely different outcome and which therefore *"...assist an (...) individual in the pursuit of her ends..."* (Sugden 1998). This cardinality equals two for the proposer in all three allocation procedures from section 2. For the responder, this cardinality equals one (or zero) in the dictator game, two in the yes-no, and four in the ultimatum game. We can now use this notion to compare the distribution of decision rights across individuals much the same way as we compare material payoffs in the inequity aversion models: there is inequality in decision rights in the dictator and the ultimatum game, yet decision rights are equally distributed in the yes-no game. It is easy to imagine that this inequality is felt more strongly if a player has lesser rights than her opponent (disadvantageous inequality), than if it is the opponent who has lesser rights (advantageous inequality for the player). More formally, for players $i = 1, 2$, let S_i be the strategy set of player i in the two-player extensive form game Γ and let α_i and β_i be the degree to which player i dislikes disadvantageous, and advantageous inequality, respectively. Then, if player i cares about the equality of decision rights and utility is linear in cardinality differences, her preferences might be characterized by:

$$u_i(s_i, s_j; b_i, b_j) - \beta_i max\{\#S_i - \#S_j, 0\} - \alpha_i max\{\#S_j - \#S_i, 0\}$$

where $u_i(s_i, s_j; b_i, b_j)$ captures that part of the social utility function which refers to selfish and other-regarding material payoffs s_i, s_j (as in Fehr and Schmidt 1999; Bolton and Ockenfels, 2000, for instance) and possibly on players' belief systems b_i, b_j (as in psychological games (Battigalli and Dufwenberg, 2009).[16] Similarly, parties may dislike unequal rights of information, an idea which we formalize in app. D where we also describe that increasing a player's decision rights comes at the cost of procedural complexity and inefficiency. Table A3 applies these criteria to all games. These formalizations are mere thought experiments to open avenues for argumentation in purely procedural terms.

5 Experimental setup

The computerized experiment was conducted in the laboratory of the Max Planck Institute of Economics in Jena. Participants were 352 undergraduates from the University of Jena, randomly drawn

[15] We use the strategy notion of an extensive form game just for formal simplicity. An individual may rather think in terms of the options laid out before her at a given decision node (her action set). For this case, one can use the sum cardinality over a player's action sets across all information sets of this player which leads to the exact same conclusions.

[16] To see that the Fehr-Schmidt (1999) model, for instance is a special case, note that $u_i(s_i, s_j) = \pi_i(s_i, s_j) - \beta max\{\pi_i(s_i, s_j) - \pi_j(s_j, s_i), 0\} - \alpha max\{\pi_j(s_j, s_j) - \pi_i(s_i, s_j), 0\}$ where $\pi_i(s_i, s_j)$ is the material payoff to player i and payoffs do not depend explicitly on beliefs. Since $u_i(s_i, s_j; b_i, b_j)$ depends on beliefs, one can also derive the payoff functions of psychological game theory – see section 3.2 – as a special case. Note, however, that adding purely procedural terms need not necessarily increase the complexity of the utility function: if a player compensates an opponent for lesser decision rights by giving that opponent more payoff, the player will behave *as if* she cared for the opponent's payout whereas she actually cares for the opponent's position of rights. Hence, only one of the two terms is needed to capture her behaviour. Yet, the distributive fairness terms would not correctly predict this player's behaviour for another game (Blanco et al. 2011) with a different distribution of rights while the purely procedural terms would.

from different fields of study. 186 of them participated in sessions which introduced the mini yes-no and the mini ultimatum game from section 2, another 166 participants in sessions which introduced the mini dictator and the mini ultimatum game. Participants were recruited using the ORSEE software (Greiner 2004). The experiment was programmed in z-Tree (Fischbacher 2007). At the beginning of each session, participants were randomly seated at visually isolated computer terminals where they received a hardcopy of the German instructions which can be found in section A.1. Subsequently, participants answered a control questionnaire to ensure their understanding, see section A.2. The experiment started after all participants had successfully completed the questionnaire. Each session introduced only one pair of procedures, either the ultimatum and yes-no game, or the dictator and the ultimatum game from section 2. In each game, a pie of 200 units (6 Euros) was to be shared. We elicited subjects' choices in all games by means of the vector strategy method (Selten 1967), that is, by asking subjects to decide in every decision node of either procedure, and for either role. We explicitly wished to exploit potential behavioural effects of the strategy vector method which familiarizes subjects with both roles and both procedures thus increasing the share of subjects who would deem the procedures outcome-invariant.[17] Subsequently, each subject in a randomly formed pair of subjects, was randomly assigned the role of a proposer or a responder.

Informed about their actual role, subjects were given an ex ante unannounced option to influence the draw of the procedure. Subjects received additional instructions on their screens to explain the option, see A.3 and answered a further control question, see A.4. Each subject then stated whether she preferred any procedure at all, and if so, which one. Subsequently, subjects could pay 15 (Euro)Cents to make their preferred procedure more likely to occur. Subjects knew that in the end of the experiment, one player in each pair would be randomly selected and her decision implemented. Subsequently, first-order beliefs were elicited. For every decision node of the opponent in either procedure, we asked a player how she believed the other player would decide. Subjects were asked how many out of four randomly drawn players of the other role they believed had made a specific choice[18]. Beliefs were incentivized such that subjects earned 100 additional units (3 Euros) for a correct answer and no additional units otherwise. This simple belief measure has recently been suggested in (Schlag and Tremewan 2012). Unlike the seemingly precise quadratic scoring rule, for instance, this simple belief measure is unbiased even for risk-averse subjects (Schlag and Tremewan 2012) and can easily, and controllably be absorbed (Price 1998). Just as the quadratic scoring rule, our simple belief measure may be imprecise. These imprecisions could bring out about challenges in our context. We discuss and dismiss those challenges in sections 6.2 and 7.1.[19]

Finally, the procedures were drawn. If the randomly selected player had stated a preference for a procedure and paid for it, then her preferred procedure was used. If she had not paid, each procedure was drawn with equal probability. The cost of influencing the procedure was subtracted. If a subject

[17]The strategy vector method does, however, not cause our results: in a recent application of this paper's preferences, Chlaß and Riener (2015) apply the play-method and find the exact same motivation underneath subjects' procedural choices over games with fair, and unfair distributions of rights – even when subjects decide only for their own role.

[18]We did not elicit subjects' beliefs about whether, and for which procedure, the other player in her pair would pay.

[19]See (Schlag and Tremewan 2012) for a comprehensive account of belief-elicitation methods.

wanted to pay but was not drawn, she would not incur any cost. Only the choices that had been made in the procedure which was drawn became payoff-relevant. To assess the correctness of a player's beliefs, four subjects of the other role were randomly drawn to see whether their behaviour coincided with the player's beliefs.

At the end of each session, we handed out a standardized moral judgement test (M-J-T) by Georg Lind (1978, 2008). The test elicits to which ethical ideals individuals actually resort in order to derive the right course of action, or put differently, how they make moral judgements. If subjects' procedural choices are motivated by *none* of the normative ideals upon which existing preference types – see section 2 – build, we can reaffirm that we report evidence for new preferences. The specific ethical (normative) ideals which subjects can employ in the test are taken from Lawrence Kohlberg's (1969, 1984) field work. Kohlberg studied extensively which ideals individuals in the field use to make moral judgements and classified them into six classes, see appendix C. Coincidentally, this classification covers all normative (ethical) ideals which economics has used to formalize preferences to date. The test presents two stories and asks subjects whether they deem the respective protagonist's behaviour right or wrong. Subsequently, the test lists an inventory of 24 arguments (12 after each story two of which refer to the same Kohlberg class in app. C) and asks subjects to agree or disagree with the use of each argument on a nine-point scale[20]. Eight arguments refer to whether the outcomes of the respective action were favourable for the self (e.g. "the action was good/right because it made me a lot of money"), another eight to the protagonist's intention, others' expectations and social norms (e.g. "the action was good/right because I intended only good outcomes for others/ was expected of me/ everybody would have done it"), and another eight refer *solely* to the way how the action came about (e.g. "the action was good, because when it was executed, others' equal rights of participation and information were respected"). Specifically the last type of argument does not refer in any way to actual, expected, or intended outcomes of this action, see section 6.2.

The test does not refer in any way to the experiment and is designed such that even in a sample of subjects who do not necessarily give their real opinion in the test (who do, for instance, try to answer in what they deem a socially appropriate and acceptable way), the true underlying distribution of the score in this sample is not biased (Wasel 1994; Lind 2002).[21] Finally, additional self-designed questionnaires were distributed in some sessions which asked subjects to also rank the procedures in terms of 'simplicity', and which gave subjects space to explain their procedural choice – if any – in written open form. We use this questionnaire data to investigate a possible existence of purely procedural simplicity, and efficiency concerns, see appendices D.3 and D.4. These concerns could be active only in

[20]This is a rating approach and therefore, subjects can rate all arguments equally should they so wish. We do not cluster subjects into 'types'. Instead, we use each subject's complete set of six scores (one preference over using each of the six Kohlberg classes) to explain her procedural choice in 6.2 and 7.2.

[21]The test achieves these desirable psychometric properties by listing arguments of the six Kohlberg classes in a different order every time (out of four times), and by varying the frames in which the four arguments pertaining to the same class are presented (two arguments out of four may be used to argue *for*, and two *against* an action) – see appendix A.5 for an excerpt. It is hence hard if not impossible for an experimental subject to see how she can fill out the test such that her score (whose calculation she does not know anyway) shifts into a specific direction. For this reason, a subject who answers specific test questions in a way to justify some or all of the many choices made during the experiment, changes her score but not into the direction she intends. Rather, she adds noise to her true score.

cases where procedural choices cannot be statistically explained by any ethical (fairness) ideal: neither by the fairness ideals underlying existing preference models in section 3, nor by a concern about the equality of rights. To avoid making any inference about a possible existence of these concerns from such a mere non-rejected Null hypothesis, we try to additionally explain subjects' procedural choices in these cases by simplicity rankings and efficiency statements.

In our analysis, we first focus on subjects who – according to *all* preference models discussed in section 3 as judged by our simple unbiased belief measure – deem their pair of procedures outcome-invariant (henceforth 'EQ'-subjects). These are responders who i) accept each proposal in each procedure and who ii) expect that the fair proposal is *always* proposed in both procedures. Proposers in turn need to i) *always* make the fair proposal and ii) think that responders *always* accept both proposals in both procedures (in the dictator game, this is satisfied by construction since responders cannot influence payoffs). We test whether these 'EQ'-subjects still prefer one procedure over the other. If this is the case, we say that subjects qualify for a *purely* procedural concern. In section 6.2, we show that these choices also identify purely procedural concerns because they can indeed be explained by the degree to which subjects resort to the equality of basic rights and liberties (Kohlbergian classes 5 or 6) in the moral judgement test, after controlling for the complete set of other ideals – the other four Kohlberg classes – they might have used.

If Kohlberg classes 5 or 6 are indeed at play, we can, first of all, rule out that 'EQ'-subjects' procedural choices are mistakes. Second, we have double checked that the existing preference models from section 3 *do not* explain the procedural choices at hand and that, consequently, the simple unbiased belief measure we use was precise enough to control for this. Third, we have provided *positive evidence* that the new preferences from section 4 are at play, and that, moreover, we did not induce an experimenter demand effect in which case subjects' choices would have been explained by their desire to comply with others' (our own) expectations about their choices (Kohlberg class 3).

Fourth, we can extend our analysis to non-'EQ'-participants by instrumenting purely procedural preferences with Kohlberg classes 5 or 6 to estimate how prevalent purely procedural preferences are overall, and which types of behaviour they induce when procedures are perceived to generate different outcomes. This is done in section 7.2.1 by building groups of non-'EQ'-subjects whose beliefs and actions differ *similarly* across procedures. All subjects *within* the same group therefore perceive a highly similar material, kindness-based, or fairness-based difference between the procedures and should therefore, make largely similar procedural choices. If instead, the procedural choices within such a highly similar group are very heterogeneous, and if this heterogeneity can be explained by Kohlberg classes 5 or 6, we have instrumental evidence for non-'EQ' subjects who are motivated by purely procedural preferences. Fifth, we can test whether 'EQ'- subjects differ persistently from other subjects in those characteristics which explain 'EQ'-subjects' purely procedural choices in 7.2. If this were true, then imposing the theoretical conditions which identify purely procedural preferences from table 2 would have led to a selection effect[22]. We do, however, not find evidence for such an effect, see section 7.2. in the next section, we start with the analysis of 'EQ'-subjects and their ethical motivations.

[22]A selection would exist if 'EQ'-subjects differed from non-'EQ' subjects – that is, subjects with procedurally *variant* actions and beliefs – in a latent characteristic which is *critical* for a purely procedural choice.

6 Results

6.1 How often do 'EQ'-subjects state a purely procedural concern?

We first concentrate on so-called 'EQ'-subjects who have actual behaviour and beliefs such that they fulfill even the most restrictive conditions for procedural indifference from section 3. These conditions ensure indifference not only in equilibrium (and would hold within such an equilibrium), but importantly, ensure indifference also off-equilibrium. 'EQ' proposers select the fair proposal in both procedures which they encounter and expect all proposals to be accepted with certainty in both procedures. 'EQ' responders accept all proposals (if the procedures allow to do so) and expect the fair proposal to be proposed with certainty in all procedures.

59% of all 'EQ'-subjects *state* a purely procedural preference, i.e. state a preference for some procedure. The 99% confidence interval for this share of 59% has lower bound of 48% and upper bound of 70%. We denote these confidence intervals by square brackets, i.e. $]48\%, 70\%[$. 21% of all 'EQ'-subjects would also pay for their preferred procedure and thereby *reveal* a purely procedural preference. The share of subjects who would pay has a 99% confidence interval of $]13\%, 32\%[$.

RESULT 1. A significant share of EQ-subjects state and is willing to pay for a *purely* procedural preference.

Looking at 'EQ'-subjects who choose between a mini dictator and a mini ultimatum game, 65% state a preference for one procedure over another. This share has a 99% confidence interval of $]47\%, 80\%[$. More specifically, 58%, $]40\%, 75\%[$, of 'EQ'-subjects who choose between the dictator and the ultimatum game state a preference for the dictator game. 25%, $]12\%, 42\%[$, of them also pay for and thereby *reveal* this preference. Only 7% within $]1\%, 20\%[$ state to prefer the ultimatum game over the dictator game and nobody, i.e. 3%, $]0\%, 15\%[$, reveals this preference.

RESULT 2. A significant share of 'EQ'-subjects states to prefer the dictator over the ultimatum game and is willing to pay for this preference.

Table 3 reviews our results for subjects who choose between the mini ultimatum (UG) and the mini dictator game (DG) along with the absolute frequency of 'EQ'-proposers, 'EQ'-responders, their respective choices, and 99% confidence intervals for the frequency of these choices.

Looking at subjects who choose between a mini yes-no, and a mini ultimatum game, 55%, $]40\%, 70\%[$ state a preference for one procedure over the other, and 16%, $]7\%, 30\%[$ *reveal* such a preference. In this pair of procedures, 'EQ'-subjects most frequently prefer the yes-no game over the ultimatum game. A share of 34%, $]20\%, 49\%[$, states this preference, and a share of 14%, $]5\%, 27\%[$ *reveals*, i.e. would pay for it. A preference for the ultimatum game over the yes-no game in turn is less frequent; only 21%, $]10\%, 36\%[$ state such a preference and only 3%, $]0\%, 12\%[$, would pay for it.

RESULT 3. A significant share of 'EQ'-subjects states to prefer the yes-no over the ultimatum game and is willing to pay for the respective preference.

role	nr. of obs.	DG $\succ$ UG		UG $\succ$ DG	
		stated	revealed	stated	revealed
proposer	35	28 $]57\%, 94\%[$[23]	10 $]11\%, 52\%[$	1 $[0\%, 20\%[$	0 $[0\%, 15\%[$
responder	25	7 $]8\%, 56\%[$	5 $]4\%, 47\%[$	3 $]1\%, 38\%[$	2 $]0\%, 33\%[$
all	60	35 (58%) $]40\%, 75\%[$	15 (25%) $]12\%, 42\%[$	4 (7%) $]1\%, 20\%[$	2 (3%) $]0\%, 15\%[$

Table 3: **39 of 60 (65%) 'EQ'-subjects state to prefer; 17 of 60 (28%) are willing to pay for a dictator or an ultimatum game** [99% **confidence intervals in brackets].**

Table 4 reviews our results for subjects' choices between a mini yes-no, and a mini ultimatum game, i.e. the number of 'EQ'-proposers, 'EQ'-responders, their respective choices, and the 99% confidence intervals for the frequency of these choices. Note that parties differ in their procedural choices and that proposers opt far more often for the yes-no game than responders do. This asymmetry and the overall popularity of the yes-no game are hard to explain in terms of outcomes: on aggregate, there are *fewer* proposers who always offer an equal split and who also expect all four responders accept this split in the yes-no game than there are such proposers in the ultimatum game, see section 7.1 for a detailed discussion. However, preferences for the equality of decision and/or information rights across parties as formulated in section 4 and appendix D.1 would predict exactly such an asymmetry in choices. As we argued there, a player i is likely to feel a stronger dislike toward disadvantageous inequality in rights, than toward advantageous inequality, i.e. $\alpha_i \geq \beta_i$, and would thus also show a stronger preference to have her own disadvantage removed than to remove the opponent's disadvantage (which is the player's own advantage). Proposers who opt for the yes-no game remove their own disadvantageous inequality in information, and in decision rights from the ultimatum game. Responders who opt for the yes-no game remove their opponent's (the proposer's) disadvantageous inequality in decision and in information rights from the ultimatum game. The preferences in section 4 would therefore predict that proposers opt more often for the yes-no game than responders.

role	nr. of obs.	UG $\succ$ YNG		YNG $\succ$ UG	
		stated	revealed	stated	revealed
proposer	42	4 $]1\%, 28\%[$	2 $[0, 21\%[$	18 $]23\%, 64\%[$	8 $]6\%, 39\%[$
responder	38	13 $]16\%, 57\%[$	0 $[0\%, 14\%[$	9 $]8\%, 46\%[$	3 $]0\%, 0.27\%]$
all	80	17 (21%) $]10\%, 36\%[$	2 (3%) $]0\%, 12\%[$	27 (34%) $]20\%, 49\%[$	11 (14%) $]5\%, 27\%[$

Table 4: **44 of 80 (55%) of 'EQ'-subjects state to prefer; 13 (16%) are willing to pay for a yes-no or an ultimatum game** [99% **confidence intervals in brackets].**

The next section explores whether and how purely procedural choices relate to individuals' conception of fairness. As pointed out in section 5, to support the claim that we identify new preferences, we need to show that the procedural choices described in this section can be explained by a new

[23] A square bracket $[57\%, 94\%]$ denotes a confidence interval which includes both end points, $]57\%, 94\%[$ denotes one which does not include either end point, $]57\%, 94\%]$ denotes one that does not include the lower end point, and $[57\%, 94\%[$ denotes one which does not include the upper end point.

fairness conception which is different from any of the ethical ideals underlying the preference models from section 3, notably different from intentionality concerns, social image, others' expectations, the absence of punishment or presence of a reward, status, and social norms. These normative ideals all lead to ultimately outcome-based preference models.

6.2 What motivates 'EQ'-subjects' purely procedural choices?

We have suggested that purely procedural preferences might reflect a desire that the rules of the game grant parties equal rights of information, and/or equal freedom of choice, see section 4 and appendix D. Such a claim for being treated equally by the rules of a game when outcomes are the same would be morally motivated ('it is unfair/immoral to favour one person over another by granting her more rights or greater privileges'). If true, an 'EQ' individual who has a purely procedural concern should put greater weight on *moral* arguments which refer to institutions and individual rights when she judges whether something is right or wrong than an 'EQ' individual who shows no purely procedural concern.

To test this, we first need a means to describe how individuals typically derive whether something is right or wrong – i.e. how they make a moral judgement – and which arguments they employ to do so. An individual typically feels comfortable to use only some of the many moral arguments which exist: each individual therefore has preferences over ways of moral argumentation (see e.g. Piaget 1948; Kohlberg 1984; Lind 2008). Kohlberg (1969, pp. 375) distinguishes three broad ways of moral argumentation: a *preconventional*, a *conventional*, and a *postconventional* way, described through altogether six 'classes of argumentation' which we review in appendix C.

An individual uses a *preconventional* argument if she argues that an action is morally right when it does not entail a punishment, or else, when that action is rewarded (classes 1 & 2). Instead, an individual uses a *conventional* moral argument, if she argues that something is morally right because it is in line with a social norm, a social expectation or done with a good intention (classes 3 & 4). Inequity aversion (Fehr and Schmidt 1999; Bolton and Ockenfels 2000; Bolton et al. 2005) and reciprocity (Dufwenberg and Kirchsteiger 2004; Falk and Fischbacher 2006) would therefore rely on conventional moral argumentation. An individual uses a *postconventional* argument if she argues that something is right because it is in line with the social contract. She may, for instance, argue that the rules of the game do not comply with the individual rights recorded in a constitution and violate the legislative principle that all individuals enjoy the same rights and that there be no discrimination (postclass 1). An individual also argues postconventionally if she deems something wrong because it violates a value or some general ethical principle which she considers universally valid, or because it infringes specific human rights *above* the social contract (postclass 2). Such a principle could be the consideration of another's will or her dignity, for instance. Specifically postconventional arguments do not refer to the outcome of an action or a process but rather, to the nature of the action or process itself. The same outcome may hence be judged very differently if generated by a dictator's decision, rather than by democratic consensus (Kohlberg 1969, p. 376).

Suppose now that the procedural choices in our experiment do indeed reflect moral preferences over

the *rules of a game* rather than preferences over outcomes, intentions, or norms. If so, there should be a strong link between individuals' procedural choices and their sensitivity to *postconventional* moral arguments. We do indeed find such a link, specifically to the first subclass of postconventional reasoning (individual rights as recorded in a constitution/the social contract). Table 5 summarizes the postconventional type of argumentation and its subclasses once more, see also appendix C.

argumentation	motivation for moral behaviour
postconventional	*postclass 1 (Kohlberg class 5).* Social contract orientation, in which duties are defined in terms of the social contract and the respect for others' rights as recorded in that contract. Emphasis is upon equality and mutual obligation within a democratic order. *postclass 2 (Kohlberg class 6).* The morality of individual principles of conscience such as the respect for the individual will, freedom of choice etc. Rightness of acts is determined by conscience in accord with comprehensive, universal and consistent ethical principles.

Table 5: **Kohlberg's two classes of postconventional moral argumentation (Ishida 2006).**

A) EQ-subjects' preferences for the mini yes-no over the mini ultimatum game. Tables 6a) and 6b) show which ways of moral argumentation among those outlined above actually link to subjects' choices of the yes-no game. Throughout, *pre*, *con*, and *post* denote the extent to which subjects make use of preconventional (Kohlberg 1 & 2), conventional (Kohlberg 3 & 4), and postconventional argumentation (Kohlberg 5 & 6), respectively. *postclass1* denotes the extent to which subjects make use of the first subclass of postconventional moral argumentation (Kohlberg 5), see table 5. Throughout, we report the marginal effect of each explanatory variable averaged over all individuals within simple binary Logit models[24].

Table 6a). **Proposers, n=42, Count R^2=0.74**

variable	*effect*	*error*	*z-stat*	*p-val.*
(pre)	(-0.09)	(0.08)	(-1.18)	(0.24)
con · post	-0.22	0.08	-2.73	0.01
postclass1	0.22	0.08	2.80	0.01

Table 6b). **Responders, n=21**[25]**, Count R^2=0.76**

variable	*effect*	*error*	*z-stat*	*p-val.*
(pre)	(-0.09)	(0.08)	(-1.13)	(0.26)
(con · post)	(-0.14)	(0.09)	(-1.56)	(0.12)
simpler[24]	0.09	0.04	2.48	0.02

Tables 6: **'EQ'-subjects' preferences for the yes-no game link to Kohlberg's class five of moral argumentation ('postclass 1' in table 5), and a simplicity argument.**

Overall, 'EQ'-subjects' use of preconventional argumentation *pre* does not significantly link to their choice of the yes-no game. Hence, we do not find evidence for an *outcome-based* motivation in terms

[24]All models begin with comprehensive specifications including all ways of moral argumentation, Georg Lind's measures for cognitive moral ability, and all possible two-way interactions between variables. These models are reduced step by step leaving out insignificant variables. Insignificant variables of theoretical interest are reported within brackets.

[25]To shed further insight into this feature, we asked subjects in some sessions to rate the procedures relative to each other in terms of simplicity. Therefore, we only have data for 21 rather than all 38 (see table 4) 'EQ'-responders.

of mere material payoffs. The effect *con* · *post* captures the interdependence between subjects' use of conventional and postconventional arguments. If this interaction is large, neither conventional, nor postconventional arguments have stand-alone value for a subject. The higher a subject scores on *con* · *post*, the *less likely* she prefers the yes-no game. Since the interaction does not increase the likelihood to prefer the ultimatum game in this pair of procedures – see *B)* below –, it shifts likelihood to being indifferent. Conventional (i.e. intention and norm-based) argumentation *con* which is at the heart of inequity-aversion and reciprocity (Kohlberg 2 & 3), does not show any impact other than through this interaction effect and therefore promotes indifference between the procedures at hand. The second individual element of the interaction effect – *post* – is also insignificant. Hence, adding the individual elements of the interaction term *con* · *post* to the regressions in tables 6 does not change either significance or nature of these results, see footnote 24.

'EQ'-proposers and 'EQ'-responders seem to differ in their reasons to opt for the yes-no game. Proposals link to *postclass 1* arguments suggesting that proposers are purely concerned about parties' rights in each procedure[26]. Responders' choices of the yes-no game link to their simplicity ratings of the procedures rather than their use of postconventional arguments. If a responder rates the yes-no game by one point (on a 7-point scale) simpler than the ultimatum game, she is an estimated 9% more likely to prefer the yes-no game, the average marginal effect being 0.09 ($z - stat: 2.48$, $p - value < 0.02$), see table 6b). If, however, we contrast responders who prefer the yes-no game only with responders who are indifferent (and leave out responders who prefer the ultimatum game which leaves us with $n = 15$ observations), this simplicity concern vanishes, and we retrieve the positive marginal effect of *postclass 1* arguments on the likelihood that responders prefer the yes-no game, i.e. (0.72, $z - stat$: 2.82, $p - value < 0.01$).

B) Preferences for the mini ultimatum over the mini yes-no game. There are mainly responders and very few proposers who state this preference. 'EQ'-proposers who prefer the ultimatum game make again more use of postconventional arguments than proposers who are indifferent – the marginal effect of *postclass 1* on preferring the ultimatum game is 0.31 ($z - stat: 2.18$, $p - value < 0.03$) with $n = 24$ observations. Preconventional argumentation and the interaction effect *con* · *post* do not show a significant impact at the 10% level. The same holds if we contrast proposers who prefer the ultimatum game with those who prefer the yes-no game: the more often proposers prefer the ultimatum to the yes-no game, the more they use *postclass 1* arguments, the effect being 0.35 ($z-stat$: 2.11, $p - value < 0.04$) with $n = 22$. For 'EQ'-responders, we obtain similar results. The more often they use *postclass 1* arguments, the more likely they are to prefer the ultimatum game to being indifferent, the effect is: 0.89 ($z - stat: 4.36$, $p - value = 0.00$) with $n = 16$. The interaction effect *con*· *post* reduces the likelihood of preferring the ultimatum game to being indifferent (effect: -1.13, $z - stat: -3.91$, $p - value = 0.00$). Preconventional arguments – i.e. 'the ultimatum game provides an option to punish which might lead to higher offers' – are not significant at the 10% level.

[26]For payment data only, we observe a similar effect of postconventional arguments. Proposers are the more likely to report a purely procedural preference, the more they make use of *postclass 1* arguments. The estimated effect is 0.19 ($z - stat: 2.08$, $p - value < 0.04$). Preconventional or conventional arguments do not show any impact.

The belief conditions derived in section 3 – see table 2 – therefore seem to be sufficient to rule out any concerns that the procedures might entail different material and pychological outcomes. At the same time, it shows that the belief measure which we implement is precise enough to control these belief conditions. In the next section, we discuss that also the distribution of 'EQ' beliefs contradicts the idea that imprecisions in the belief measure drive the main part of the evidence, or risk aversion[27].

c) Preferences for the dictator over the mini ultimatum game. Subjects' choice of the dictator game is not significantly linked to *any* way of moral argumentation – and does therefore also not reflect a desire on the proposer side to avoid punishment for the equal split (preconventional argumentation) when proposers could have offered a generous split. Preferences for the dictator game are particularly puzzling for 'EQ'-responders who remove their own veto. The procedural property which we predict subjects to value is the procedure's *efficiency* in regulating conflict when the responder's interests are always protected: a procedural property without ethical content, see appendix D.4. To back this claim beyond a missing correlation with subjects' moral judgement, we classified subjects' answers from our open-form post experimental questionnaire where available. If a subject stated that she preferred the dictator game because '*neither* party could get a zero payoff', we classified her as efficiency concerned[28]. In other words, a procedure is efficient if *it does not allow for the destruction of the pie even when it is known that this destruction never takes place.* In the dictator game, there are only efficient strategies since neither the responder nor the proposer has an inefficient strategy as defined in Dufwenberg and Kirchsteiger (2004, pp. 276). Overall, 'EQ'-proposers who state an efficiency concern are 17% more likely to prefer the dictator game (effect: 0.17, $z - stat : 2.17$, $p - value < 0.03$) on $n = 35$[29], considering only proposers who are willing to pay for the dictator game, the effect becomes 0.33 ($z - stat : 1.97$, $p - value < 0.05$)[30] on $n = 16$ observations. For 'EQ'-responders, we obtain a similar result. Responders who state an efficiency concern are 33% more likely to prefer the dictator game (effect: 0.33, $z - stat : 3.13$, $p - value < 0.01$) with $n = 16$. For the seven responders who are willing to pay for their preferred procedures, the efficiency concern perfectly explains the preference for the dictator game compared with eight responders who do not pay. 'EQ'-responders' efficiency concern nicely illustrates the purely procedural nature of the concern. 'EQ'-responders always accept and therefore, know for sure the pie will not be destroyed. Still, they choose the dictator over the ultimatum game which removes inefficient options from their own action set which they never use.

[27] In line with the next section, we can also dismiss here that a latent correlation between risk preferences and *postclass 1* arguments explains the significant effect of *postclass 1* arguments on choices: for data on 285 subjects collected from the same subject pool at the same time, a linear (or indeed, any nonparametric) association between risk preferences *and postclass 1* arguments remain insignificant (effect: -0.03, $p - value = 0.27$)

[28] Subjects who state an efficiency concern amongst several concerns are not classified to hold a concern for 'efficiency'.

[29] As mentioned before, we do not have such statements from all participants which is why the number of proposers and responders in this paragraph are not equal to the shares of 'EQ' proposers and responders in section 6.1.

[30] On the overall set of 'EQ'-Proposers, preconventional argumentation also shows a weak effect (0.12, $z - stat : 1.87$, $p - value = 0.07$) which vanishes (0.15, $z - stat : 1.51$, $p - value = 0.13$) if we consider payment data only. Hence, proposers who merely state a preference for the dictator game acknowledge that the lack of responder veto might entail a material advantage – but not for this paper's payoffs – hence, they state to be indifferent.

7 Discussion: Robustness & Prevalence

7.1 'EQ' subjects' – precision of beliefs

If our belief measure were infinitely precise, all preference theories to date – see section 3 – would predict 'EQ'-subjects to be indifferent between procedures. Yet, infinitely precise belief measures do not exist. The quadratic scoring rule, for instance, derives its precision from the assumption that subjects are risk-neutral, an assumption which itself can only be imperfectly controlled in practice. For risk-averse EUT individuals, for instance, (Schlag et al. 2014), the seemingly precise belief reports under a quadratic scoring rule are biased away from the extremes. Apart from this bias, the method is known to be unintuitive to subjects and therefore, known to be only partly absorbed (Price 1998). In view of these problems, we chose an unbiased frequentist method which coincides with the simple belief elicitation recommended in (Schlag and Tremewan 2012):

We elicited beliefs by asking subjects how many out of 4 randomly drawn opponents choose a given one of the two available actions. Those subjects who report beliefs 4 out of 4 or 0 out of 4 are of particular interest to us since they contribute to the set of 'EQ'-subjects which is the starting point of our analysis. This belief elicitation method is intuitive and unbiased, and can easily and controllably be absorbed (Price 1998) which is not true for methods where subjects must report probability estimates. Its only disadvantage is that it does not provide point belief estimates in probabilistic terms as, for instance, the quadratic scoring rule, does (leading to an illusion of precision). Having dismissed measurement bias (see above), we are hence left with measurement imprecision. In this section, we show that we can identify purely procedural preferences despite this measurement imprecision of our frequentist belief elicitation method. The distribution of beliefs across procedures contradicts the idea that measurement imprecisions underlie the choices in section 6.1.

How imprecise is the unbiased belief elicitation method we apply? Theoretically – see Schlag and Tremewan (2012) for details – a subject who submits a belief that 4 out of 4 opponents choose a specific action, has a probabilistic confidence of 80 % or higher that all opponents choose this action. A subject in turn who submits that 0 out of 4 opponents choose that action have a probabilistic confidence of 80 % or higher that no opponent chooses this action. A concern about measurement imprecision in our case can hence be exemplified as follows: $\geq$ 80 % confidence is not equal to 100% confidence, and yet our identification method for *purely procedural preferences* requires that we identify subjects who are 100% confident that each procedure generates the same outcomes, and who still pay for a(ny) game. An argument against our claim that we find evidence for new, purely procedural preferences goes as follows: *"The majority of 'EQ' subjects prefers the yes-no game. An 'EQ' proposer who chooses between the mini-ultimatum game and the mini-yes-no-game could report a belief that 4 out of 4 responders accept the equal split in both games, and that four out of four responders also accept the generous split (20 ECUs for the proposer and 180 ECUs for the responder) in both games. Yet, this proposer might actually believe that the proposal in the yes-no-game will be accepted with probability 99 % and the fair fifty-fifty proposal in the ultimatum game with 81 % probability. If this proposer offers the equal split in both games, she would be* 0.18×100 *ECUs better*

off in the yes-no-game. Since the proposer can only influence the draw of the procedures with 50% *probability in her pair and only if she pays* 5 *ECU, she would gain* $0.5 \times (18-5) = 6.5$ *ECU by paying for the yes-no game. Therefore, this proposer's so called purely procedural preference exhibits nothing but self-interest after all."*

Let us look at the relevant set of proposers who always offer the equal split. If the counter-argument above were true, then we must – firstly – observe that there are *more* such equal split proposers who report a *4/4* acceptance belief for the yes-no game than who report such a *4/4* belief for the ultimatum game.[31] This is, however, not true: there are *less* (64) proposers who always offer the equal split who report a *4/4* belief in the yes-no game than in the ultimatum game (66).[32] Summing up, we find that – if anything – proposers and responders would each expect to hold a small material *dis*advantage in the yes-no game. Self-interest can therefore, not explain the aggregate preference for the yes-no game which was also 'EQ' subjects' main preference in 6.1[33]. Outcome-based equity theories do not explain the preference for the yes-no game either given the belief patterns mentioned: players can achieve an invisibly higher degree of expected equity by opting for the ultimatum game. Reciprocity explanations work into the same direction: if anything, the overall belief patterns suggest that both responders and proposers (with identical actions) would expect a higher psychological payoff in the ultimatum game. Hence, if parties had reciprocal preferences, they should unanimously prefer the kinder, the ultimatum game. Nevertheless, most prefer the yes-no game.

If the counter-argument were true, we should – secondly – observe that proposer choices for the yes-no-game correlate with moral argumentation from Kohlberg classes one to four, see appendix C – where material benefits, costs, social comparisons and norms, expectations and status determine what a subject deems to be the *right* course of action. This is, however, not what we observe. The evidence for purely procedural preferences in 6.1 correlated with Kohlberg class five in 6.2, a new ethical ideal upon which none of the existing preferences in section 3 builds, and an ethical ideal which explicitly refers to the equality of rights. It is also noteworthy that given the actual distribution of 'EQ' beliefs, risk-aversion would – if anything – predict that 'EQ'-subjects hold an aggregate preference for the ultimatum game where they would expect a weakly higher payoff at a lower risk.

Summing up, our evidence is indeed in line with purely procedural fairness, and at odds with outcome-based explanations building upon immeasurable differences in beliefs across games, or risk preferences. In particular, we need not make an equilibrium assumption at any point to show this. Finally, if the counter-argument were true, we should certainly not observe proposers who – motivated by the same new ethical ideal about the equality of rights – avoid the yes-no game when they expect

[31] If in the yes-no game, the acceptance likelihood were 99% and in the ultimatum game only 81%, then on a set of 84 proposers who always offer the equal split, we should observe $(0.99 - 0.81) \cdot 84 = 15$ *more* proposers with *4/4* beliefs in the yes-no game than we observe *4/4* beliefs in the ultimatum game.

[32] For 'EQ' responders, we can also reject the argument that they might in general expect an immeasurable material advantage in the yes-no game. Of 74 responders who accept all proposals in all games, 52 believe all four proposers offer the equal split in the yes-no game whereas only 47 think this is true in the ultimatum game. There are hence *more* responders who always accept and who expect all four proposers offer the generous split in the ultimatum game than there are such responders in the yes-no game.

[33] Coincidentally, the yes-no game is also the preferred according to a purely procedural preference for the equality of decision rights, see 4, the equality of information D.1, and purely procedural simplicity D.3.

a measurable material advantage (and hence, a disadvantage for the responder) for this game, but opt into this game when it does not hurt the responder and hence, is to their own disadvantage. Yet, sections 7.1-7.3 assemble these pieces of evidence which allow us to brush off concerns for hidden differences in beliefs and explore the robustness of our findings.

dictator vs. ultimatum game. On the relevant set of proposers – those who state an efficiency concern and who always offer the equal split – 95% report a *4/4* acceptance belief for the equal split in the ultimatum game but only 63% also report such a belief for the generous proposal (which they do not offer). In the dictator game, the expected acceptance probability is by construction 100%. Given these belief patterns, the main difference between both games would therefore lie in the greater *unkindness* of the ultimatum game, if immeasurable belief differences mattered at all. Yet, we do not observe that dictator game choices link to moral argumentation underlying reciprocal preferences according to which intentions, social norms, punishment avoidance, or a material cost-benefit analysis (Kohlberg classes one to three) determine the right course of action. On the responder side, the 65 who always accept report altogether more *4/4* equal split beliefs for the dictator than for the ultimatum game which implies a payoff advantage in the ultimatum game. Hence, self-interest or risk aversion could not explain why 'EQ' responders prefer the dictator game. Fairness and equity norms might be at play but in this case, responder choices of the dictator game would need to correlate with Kohlberg class three. Since i) choices of the dictator game do not correlate with any Kohlberg class, since ii) correlate with an efficiency concern, and since iii) self-interest cannot be at play given these beliefs, our evidence is again more in line with a purely procedural efficiency.

7.2 Do these motives carry over to other sets of subjects?

'EQ'-subjects' procedural choices do indeed link to normative ideals about the equality of rights (and a procedural efficiency concern) as elicited through a formal psychological test which did not relate to the incentivized experiment in any way. Hence, the outcome-based confounds from section 3 do not seem to matter which is also confirmed by the distribution of 'EQ'-beliefs across games, see section 7.1. Reasoning in terms of a normative ideal about equal rights clearly differs from reasoning in terms of expectations or intentions, as postulated by inequity aversion, guilt aversion, reciprocity models, or still, existing procedural preference models. In our setup, we therefore identify a new normative ideal which does not underlie any existing preference type. How about other subjects? Might even subjects with procedurally varying beliefs choose a specific procedure because they are concerned with individual rights or efficiency rather than with the (subjective) outcomes of that procedure?

To test this, we cluster (group)[34] all non 'EQ'- subjects according to their actions and beliefs and analyze whether their behaviour links to the same normative ideal as the procedural choices of 'EQ'-subjects did. Within each cluster, all subjects – those who do report a procedural preference, and those who report to be indifferent – have similar material, intention-, or norm-based incentives to

[34]Clusters were obtained using Ward's method; cluster similarity was measured by Eucledian distance in five dimensions: i) proposers' offer in the yes-no game, and ii) in the ultimatum game, iii) proposers' belief about how many responders accept in the yes-no game, and iv) how many responders accept the equal split and v) the generous split in the ultimatum game, respectively. Initially, each procedure produced three clusters.

choose one procedure over another. In statistical terms, we 'balance' subjects on the degree ('balancing score') to which inference about purely procedural motives from their experimental choices could theoretically be confounded (Rosenbaum and Rubin 1983) by any of the outcome-based preferences discussed in section 3. Keeping this potential confound constant across subjects who state indifference between two procedures and subjects who state a preference, we remove the strategic confound from the comparison of these groups. This method is more precise than directly estimating (Heckman 1979) the confound since it reduces specification uncertainty: we need not specify, and therefore, not be correct about how much the perceived strategic incentive changes [35] if, say, a subject believes three responders accept in the yes-no compared with only two in the ultimatum game etc. If the clustering achieves to make the strategic confound similar enough between all groups, then choosing any game over being indifferent should link to the same normative ideal as for 'EQ' subjects in 6.2. If after clustering, the strategic confound still differs between subjects who prefer different games or report indifference, we can identify subjects who choose *against* their incentive and see whether the normative ideal underlying the new preferences from section 6.2 explains subjects' willingness to forego payoff. The respective test hypotheses will then differ across clusters and a multinomial model be inadequate since each category differs from the reference category by a different strategic confound.

We now analyze whether individuals who choose a procedure *and* have a strategic incentive to do so, respond to this strategic incentive, or whether – just as their 'EQ'-counterparts – they are concerned about individual rights (or efficiency) and just coincidentally happen to believe that the procedures also generate different (subjective) outcomes. Similarly, we can test more generally whether individuals who prefer not to pursue their strategic gain (who for instance, states indifference when one game clearly yields them more payoff) do so out of a concern about the distribution of rights, or a concern about procedural efficiency, respectively.

A) Proposers with procedurally variant actions and beliefs, yes-no vs. ultimatum game. The WARD-clustering procedure on non-EQ proposers generated one cluster with #22, one with #9, and one with #20 proposers. The second cluster being too small to be analyzed, we manually merged it with cluster 1 thus keeping cluster 3 at maximal homogeneity[36]. In this merged cluster with #31 observations, proposers believe to have a material advantage in the ultimatum game, see table A4 in appendix F for details on all clusters. Those who opt for the yes-no game and decide *against* their incentive make more use of *postclass 1* arguments than those who are indifferent (effect: 0.24, $z-stat:$ 3.94, $p-value = 0.00$) with $n = 25$. Interestingly, also those proposers who act in line with their incentive and opt for the ultimatum game make more use of *postclass 1* arguments than those who are indifferent (effect: 0.29, $z-stat:$ 3.33, $p-value < 0.01$) on $n = 16$. Altogether, 15/31 (48%) of all proposers in the merged cluster prefer the yes-no game, and 6/31 (19%) prefer the ultimatum

[35]The balancing score keeps all variables between two groups identical which contribute to a potential strategic confound. Hence, all possible functional forms of the strategic confound must also be identical between the two groups – since the functions are evaluated at identical arguments.

[36]Since the results on cluster 1 before and after merging it with cluster 2 are the same, the additional heterogeneity introduced into cluster 1 is not critical. Note that only manually merging both clusters at this stage allow us to keep cluster 3 at maximal homogeneity and therefore, at maximal similarity in the strategic confound. Generating two clusters from the outset would have introduced more heterogeneity into *all* clusters and should therefore be avoided.

game. In cluster 3 with $n = 20$, 10 proposers prefer the yes-no game, and 9 state to be indifferent. Most proposers who prefer the yes-no game expect a material advantage in this game. Instead, most proposers who state to be indifferent expect a material advantage in the ultimatum game but decide *not to pursue* this advantage. These proposers make more use of *postclass 1* arguments than those who prefer the yes-no game. If we exclude the only three proposers who state to be indifferent and have yet another incentive structure, the effect turns from weak (-0.25, $z - stat: -1.98$, $p - value < 0.047$) on $n = 20$ to intermediate significance (-0.29, $z - stat: -2.38$, $p - value < 0.017$) on $n = 17$. These proposers who state indifference and at the same time expect an advantage of an average 40 ECU in the ultimatum game might not wish to materially profit from amending the transparency of the procedure by choosing the ultimatum game – see appendix D.2 for a formulation of this property.

B) Responders with procedurally variant actions and beliefs, yes-no vs. ultimatum game. The initial clusters contained #22, #21, and #12 observations, respectively. In cluster 1, responders expect a payoff advantage in the ultimatum game. Those who nevertheless prefer the yes-no game make more use of *postclass 1* arguments than responders who prefer the ultimatum game (effect: 0.46, $z - stat: 2.96$, $p - value < 0.01$) with $n = 15$. Responders who opt for the yes-no game expect to forego an average strategic advantage of 98.33 ECU. Even responders who state to be indifferent and thus do not actively pursue their average advantage of 9.28 ECU in the ultimatum game care weakly more for *postclass 1* arguments than other responders who – in line with their material incentive – opt for the ultimatum game (effect: $+0.27$, $z-stat: +1.75$, $p-value = 0.08$) with $n = 16$. Moving to cluster 2 and 3, responders believe they have a payoff advantage in the yes-no game. Responders who state to be indifferent – and hence, prefer not to actively pursue an expected average strategic advantage of 32.08 ECU – make more use of *postclass 1* arguments than those who exploit their advantage and opt for the yes-no game. We merge both clusters to obtain a reliable sample size, and find a marginal effect of *postclass 1* arguments on the likelihood of being indifferent of 0.31 ($z-stat: 4.12$, $p-value < 0.01$) with $n = 21$. Responders who prefer stating indifference over opting for the ultimatum game, make more use of *postclass 1* arguments, too (effect: $+0.22$, $z-stat: +2.15$, $p-value = 0.04$) with $n = 24$.

C) Subjects with procedurally variant beliefs, dictator vs. ultimatum game. Stated efficiency concerns perfectly predict proposers' choices of the dictator game in all clusters. #6 of #24 proposers choose the dictator game and state an efficiency concern in cluster 1, see also table A5 in appendix F. These efficiency-minded proposers expect a greater advantage (on average, 44.17 ECU) in the dictator game than their non-efficiency minded counterparts (24.67 ECU). Yet, only 1 efficiency-minded proposer pays for this game while 8 (of 15) non-efficiency minded proposers do so. Again, proposers who value procedural efficiency might not wish to amend this property at the material expense of the recipient. In clusters 2 and 3, we observe an analogous effect. In cluster 2, #7 of #24 proposers who opt for the dictator game and state an efficiency reason expect a material advantage in the *ultimatum game* of an average 9.29 ECU. Amending the efficiency of the game does therefore not cause any material disadvantage to the recipient. Now, nearly all (#6 out of #7) efficiency-minded proposers pay for the dictator game. Non-efficiency minded proposers expect an advantage in the dictator game of an average 11.67 ECU but only #5 out of #12 of them pay for it. Altogether, 'efficiency' statements explain the dictator game choices for 27% of all non 'EQ'-proposers within a 99% confidence interval

of [12%, 47%]. Turning to responders, stated efficiency concerns perfectly predict responder choices of the dictator game in all clusters. In cluster 1, #6 out of #33 responders (18%) choose the dictator game and state an efficiency concern. #5 of #6 efficiency-minded responders pay while expecting, on average, a material advantage of only 10 ECU. Non-efficiency minded responders expect three times this advantage (30.94 ECU) in the dictator game but only #9 of #16 pay for it. In cluster 2, #3 of #13 responders (23%) state an efficiency concern and choose the dictator game while expecting a material *dis*advantage of 20 ECU. #2 of #12 (17%)responders do so in cluster 3 expecting a material disadvantage of 5 ECU while non-efficiency minded counterparts expect an average advantage of 20 ECU. Altogether, responders who choose the dictator game for its 'efficiency' account for 19% of all non 'EQ'-responders with a 99% confidence interval of [8%, 36%]).

Table 7 shows postestimation results for each of the clusters above. We identify the critical threshold of *postclass 1* arguments for which the predicted outcome in a given Logit model changes and report the number of participants who score above this critical threshold. For choices between the dictator and ultimatum game, this amounts to counting who states an efficiency concern and opts for the dictator game since these correlate perfectly. Altogether, we obtain the estimated shares of non 'EQ' participants who act out of the same purely procedural motivation as 'EQ'-subjects did in section 6.2 which extends the analysis from sections 6.1 and 6.2 to the full set of participants.

role	cluster (#nr. of obs.)	UG vs. YNG	
proposer	cluster 1 (#31)	21 (68%)	[43%, 87%]
	cluster 2&3 (# 17+#3)	6 (30%)	[8%, 61%]
responder	cluster 1 (# 22)	6[37](27%)	[7%, 57%]
	cluster 2&3 (# 33)	18 (55%)	[31%, 76%]
all	106	51 (48%)	[35%, 61%]

role	nr. of obs.	DG $\succ$ UG, DG vs. UG	
proposer	cluster 1 (# 24)	6 (25%)	[7%, 53%]
	cluster 2 (# 24)	7 (29%)	[9%, 58%]
responder	cluster 1 (# 33)	6 (25%)	[7%, 53%]
	cluster 2 (# 13)	3 (23%)	[3%, 62%]
	cluster 3 (# 12)	2 (17%)	]0%, 58%]
all	106	24 (23%)	[13%, 35%]

Tables 7: **Posterior Logit estimates indicate 48% of non 'EQ' subjects change their behaviour for *postclass 1* arguments (left table), 23% for efficiency arguments (right table).**

To sum up, we find that the new normative ideal is at play in all sets of procedurally varying beliefs and behaviour and hence, amongst all types of non-'EQ' subjects. On the one hand, there are subjects who still choose a given procedure due to *postclass 1* arguments or purely procedural efficiency concerns even in the presence of a small material confound. In these cases, the material confound which we measure is either too small to crowd out this new ethical ideal, or the material incentive is too small to be perceived. On the other hand, the new ethical ideal underneath this paper's purely procedural preferences – see 6.2 – also explains statistically why many subjects choose *against* their incentives. Interestingly, the interaction effect $con \cdot post$ which *reduced* the likelihood of a purely procedural concern on the set of 'EQ'-subjects is never significant for non 'EQ'-subjects. Purely procedural concerns might hence be *more* frequent among non 'EQ'- than among 'EQ'-subjects.

[37]We use only Logits where *postclass 1* arguments had a marginal effect with $p - value < 0.05$. If we also consider weaker significance levels, there are further estimated 5 responders in cluster 1 who change their behaviour out of a *postclass 1* motivation. These responders expect a payoff advantage in the ultimatum game but state to be indifferent.

7.3 Is there a selection effect?

Section 6 first concentrated on 'EQ'-subjects who should – even according to the most restrictive conditions from section 3 – be indifferent between procedures. These subjects should have no outcome-based, intention-based, or expectation-based motive to prefer one procedure over another. To understand the nature of 'EQ'-subjects' procedural choices, we studied whether and how these choices related to the moral ideals which subjects employed to determine whether a course of action is morally right or wrong. 'EQ'-subjects more likely preferred one procedure over another, the more often they argued in terms of individual rights (e.g of information and participation) as stipulated by a social contract when making such a moral judgement. There did, therefore, indeed seem to be a moral ideal at play which was outcome-independent as we require. We also observed purely procedural choices of 'EQ'-subjects which did not link to individuals' moral judgement but linked to subjects' simplicity rankings of the procedures, or to self-reported concerns for purely procedural efficiency.

Do the procedural choices which we reported for 'EQ'-subjects in section 6 result from a selection effect? A selection effect would imply that 'EQ'-subjects differ from all other subjects in some characteristic which is *critical* for a purely procedural choice, and that therefore, the new type of preference which we report is either significantly more, or less prevalent, in non 'EQ'- than in 'EQ'-subjects. To test for such an effect, we use the motivations behind 'EQ'-subjects' purely procedural choices – the characteristics which were *critical* for their purely procedural choices – and test whether these motivations are per se more relevant to 'EQ'-, than to non 'EQ'-subjects.[38]

Moral argumentation & simplicity. We could not confirm that 'EQ'-proposers or 'EQ'-responders differ from their non-'EQ' counterparts when making a moral judgement. Specifically, 'EQ'-proposers and 'EQ'-responders cannot be confirmed to make *more* use of those moral arguments – i.e. the first class of postconventional arguments *postclass1*, see section 6.2 – which were positively linked to the purely procedural choices we report (Wilcoxon Rank Sum tests, proposers: $p - value = 0.67$, responders: $p - value = 0.60$). Moreover, 'EQ'-proposers and 'EQ'-responders cannot be confirmed to score *lower* on variable $con \cdot post$ which was negatively linked to purely procedural choices and which therefore makes these choices *less* likely (Wilcoxon Rank Sum tests, proposers: $p - value = 0.62$, responders: $p - value = 0.40$). Comparing the simplicity rankings, 'EQ'-responders deem the yes-no game less often simpler than the ultimatum game than non-'EQ' responders (exact Wilcoxon Rank Sum test, $p - value < 0.05$). *A negative selection effect* might therefore have occurred in section 6.1 by underestimating the frequency of responders preferring the yes-no game.

For each motive, we also derive the critical 'strength' at which the binary logit models in section 6.2 start to predict a purely procedural choice, if all other explanatory variables take on their mean value and perform Fisher's exact test to see whether there are significantly more 'EQ'-, than non-'EQ'-subjects who score above this critical threshold. We did not find any significant difference for any explanatory variable in any type of procedural choice, or any role. 'EQ'- and non 'EQ'-responders

[38]The selection effect could also operate such that a link between these motivations and a purely procedural preference exists exclusively in 'EQ'-subjects. However, we have shown in the previous section that this is not the case.

do not even differ in their simplicity rankings of the procedures around the respective critical threshold. However, the 45% of proposers who care *most* for postconventional argumentation always have non-'EQ' beliefs and actions. Some proposers might choose procedurally variant actions or hold procedurally variant beliefs *because* they deem the procedures unjust.

Efficiency motive. Many 'EQ'-proposers and responders preferring the dictator over the ultimatum game stated in an open form post experimental questionnaire that they did so because the dictator game prevents zero payoffs for *either* party. The purely procedural nature of this efficiency concern was particularly credible for 'EQ' responders: knowing that they would always accept in both games, and expecting the equal split for sure, they opted for the procedure where they had no influence at all. While 45% of all 'EQ'-subjects ('EQ'-proposers: 39%, 'EQ'-responders: 58%) stated this reason for their choice, also 33% of all non 'EQ'-subjects (proposers: 33%, responders: 33%) did so. This is surprising since for these belief conditions, one would have expected either self-interest, or an outcome based other-regarding concern to matter. Again, the efficiency motive is not reported significantly more often by either 'EQ'-proposers or 'EQ'-responders than by their non 'EQ'-counterparts.

8 Conclusion

We present evidence that agents care about procedures in a way which no existing preference model can explain: agents prefer certain procedures over others even when they do not expect these preferred procedures to generate more advantageous, more equal, or kinder outcomes. Moving away from these conditions which allow us to infer the existence of the preferences we suspect, we show that such preferences produce a wide range of behavioural anomalies amongst our participants.

Procedural preferences are not new, *purely procedural preferences* are[39]. So far, economists interested in procedural concerns have focused on preferences for fair randomizations over unequal outcomes, e.g. (Bolton et al. 2005), or preferences for procedures which generate *kind* distributions of outcomes (Sebald 2010). In both approaches, procedural preferences are conceived as preferences over the outcomes which different procedures generate. Even in psychology – a field with a long-standing empirical interest in procedural justice – procedural preferences have an inevitable distributive foundation, see (Cropanzano and Ambrose 2001) for an extensive review and a discussion of the psychological literature on 'purely' procedural preferences.

In this paper, however, we find evidence for preferences that a procedure itself should meet certain criteria which do not refer to the distribution of outcomes generated by this procedure at all: that it should distribute decision and information rights equally across parties, that it should be simple, and efficient. We also report instances where these normative ideals explain individuals' decision to forego strategic advantages in a given allocation procedure.

[39]The idea that the rules of the game by themselves may affect utility, is not new, see e.g. (Benz and Stutzer 2003). In a survey study, Frey and Stutzer (2005) report that self-reported happiness increases in citizens' democratic rights. Yet, this self-reported happiness can be the result of those procedural preferences ultimately based on outcomes which we discussed in section 3, of an improvement in citizens' life circumstances (the outcomes of the political process), as well as the right of participation (in the political process) itself. We study the existence of purely procedural preferences in a controlled setting and also find a concern for efficiency which may moderate concerns for increased participation.

In two-player pie-sharing procedures which yield the same expected material and psychological/other-regarding *equilibrium* payoffs, we find that subjects who – according to all social preference theories known to date – should be indifferent *in and off equilibrium*, still show preferences over the procedures at hand. Subjects therefore seem to care for *purely procedural criteria* – or put differently, they seem to care for *the rules of the game* without any reference to outcomes.

We provide supporting evidence that there are outcome-invariant moral ideals behind these purely procedural concerns. Scores from a standardized moral judgement test (Lind 1978; Lind 2000; Lind 2008) measuring individuals' preferred ways of moral argumentation consistently predict subjects' preferences for a procedure.[40] The more subjects use arguments which refer to the respect for individual rights stipulated by the social contract when making a moral judgement, the higher is the estimated likelihood to prefer one pie-sharing procedure over the other (when behavioural theories unanimously claim subjects should be indifferent *in and off equilibrium*). We use this result we extend our analysis to the entire set of subjects. Many procedural choices by subjects who ignore or actively forego expected payoff advantages can be modelled by exactly the same ethical ideals as on the set of indifferent subjects. This holds for 48% (51 out of 106) of those participants who should not be indifferent between the yes-no and the ultimatum game, and for 23% of those participants who should not be indifferent between the dictator and the ultimatum game (24 out of 106). Some subjects seemingly try to 'compensate' the rules of the game behaviourally.

Why care about purely procedural preferences? One might argue that the core interest of the economic discipline lies in observed choices and outcomes, and neither in the personal nor in the institutional decision making procedures behind these (Gul and Pesendorfer 2005). Yet, take the election example from the introduction again. More individuals may vote in a procedurally fair than in an unfair election, even if the same candidate is expected to win equally likely in both cases. High abstention rates may undermine a democratic process by reducing the legitimacy of the winning candidate and trigger institutional change in the long run. Moreover, voters who find that electoral rules violate their moral ideals – e.g. by giving some minority less rights to participate, or less information – may change their votes in the interest of that minority to compensate the infringement of the minority's rights. This would be one example how individuals who respect individual rights and the social contract may compensate the rules of the game by *altruistic* behaviour.

Consistent with this idea, Chlaß and Moffatt (2012) find that dictators' propensity to give in standard anonymous dictator games strongly increases in dictators' value of universal individual rights.[41] Notice that a dictator game is a procedure which denies the recipient any right to state her own will. There is also evidence that individuals show a distaste for discriminatory taxes, even when they are socially as efficient as non-discriminatory taxes and produce the same expected outcomes (Tyran and Sausgruber 2014). Purely procedural concerns may therefore also have implications for fiscal policy.

[40]Sociologist Jean Piaget and psychologist Lawrence Kohlberg studied the types of moral argumentation which individuals in the field use when making a moral judgement. In Lind's (1978) test, subjects are asked to moral judgements using precisely these types of argumentation reported by Piaget and Kohlberg. Subjects' preferences over pie-sharing procedures in our experiment link to the extent to which they invoke individual rights and the social contract in their moral judgement – exactly the normative ideal underlying the preferences we formulate.

[41]This finding is robust under different frames, and under real-effort conditions.

Chlaß and Riener (2015) use this paper's preferences to explain when individuals are averse against sabotaging an opponent, and averse against telling selfish black lies.

Similarly, purely procedural preferences could explain why, across simple pie-sharing games and across different roles in one game, the same individual shows behaviour which is largely inconsistent with stable inequity-averse preferences (Blanco et al. 2011). One can compensate an opponent for the unfair rules of one game by being altruistic, and behave fully selfishly under fairer rules without being inconsistent. Indeed, Shor (2009) finds that when proposers are first allowed to choose between a dictator and an ultimatum game, they rather choose the ultimatum game. Moreover, those who do choose the dictator game give more than those who choose the ultimatum game; they also give more than dictators who could not choose any procedure. Shor coins this an "innate preference for voice, a key component of procedural fairness." Our paper does yet also show that individuals have a competing need for procedural efficiency and remove their own veto – their own freedom of choice – to achieve it provided their right to express an opinion is left intact.[42]

[42]Neri and Rommelsdorfer (2014) suggest that individuals prefer procedures where nobody can interfere with their decisions. In our paper, efficiency-minded individuals remove their own right to interfere. These subjects do therefore also seem to care for procedural efficiency. Dana et al. (2007) show, however, that individuals may try to prevent the opponent from learning the rules of the game – i.e. from knowing that a pie *can* actually be shared – knowledge which is crucial for the opponent to care about efficiency.

References

Adams, J.S. (1965), Inequity in Social Exchange, Berkowitz, L. (Ed.), Advances in Experimental Social Psychology, Academic Press, NY.

Albrecht, M., Mertins, V. (2015) Does Participation Increase Outcome-Acceptance?, Journal of Economics and Statistics, Issue 2.

Aldashev, G., Kirchsteiger, G., Sebald, A. (2010), How (Not) to Decide: Procedural Games, Manuscript, Université Libre de Bruxelles.

Bartling, B., Fehr, E., Herz, H. (2014) The Intrinsic Value of Decision Rights, Econometrica, 82(6), pp. 2005-2039.

Battigalli, P. and M. Dufwenberg (2007), Guilt in Games, American Economic Review, Papers and Proceedings, 97, pp. 170-176.

Benz, M., Stutzer, A. (2003), Do Workers Enjoy Procedural Utility?, Applied Economics Quarterly, 49, pp. 149-172.

Bicchieri, C. (2006). The Grammar of Society, Cambridge, UK: Cambridge University Press.

Binmore, K. (1994). Game Theory and the Social Contract, Vol. 1: Playing Fair, MIT Press.

Blanco, M., Engelmann, D., Normann, H. T. (2011), A Within-Subject Analysis of Other-Regarding Preferences, Games and Economic Behavior, 72, pp. 321-338.

Blount, S. (1995), When Social Outcomes Arent Fair: The Effect of Causal Attributions on Preferences, Organizational Behavior and Human Decision Processes, 63, pp. 131-144.

Bolton, G. (1991), A Comparative Model of Bargaining: Theory and Evidence, American Economic Review, 81, pp. 1096-1136.

Bolton, G., Ockenfels A. (2000), ERC - A Theory of Equity, Reciprocity and Competition, American Economic Review, 90, pp. 166-193.

Bolton, G., Brandts, J., Ockenfels A. (2005), Fair Procedures: Evidence From Games Involving Lotteries, Economic Journal, 115, pp. 1054-1076.

Brandts, J., Güth, W., Stiehler, A. (2006), I want YOU! – An Experiment Studying the Selection Effect When Assigning Distributive Power, Labor Economics, 13(1), pp. 1-17.

Charness, G. and M. Dufwenberg (2006), Promises and Partnership, Econometrica, 74, pp. 1579-1601.

Charness, G., Cobo-Reyes, R., Jiménez, N., Lacomba, J. A., Lagos, F. (2012), The Hidden Advantage of Delegation, Pareto Improvements in a Gift Exchange Game, American Economic Review 2012, 102(5), pp. 2358-2379.

Chlaß, N., Moffatt, P. (2012), Giving in Dictator Games, Experimenter Demand Effect, or Preference over the Rules of the Game?, Jena Economic Research Papers, 2012-44.

Chlaß, N., Riener, G. (2015), Spying, Lying, Sabotaging: Procedures and Consequences, University of Mannheim Working Paper #ECON 15-17 and DICE Working Paper 196-2015.

Cox, J.C., Friedman, D., Gjerstad, S. (2007), A Tractable Model of Reciprocity and Fairness, Games and Economic Behavior, 59, pp. 17-45.

Cox, J.C., Friedman, D., Sadiraj, V. (2008), Revealed Altruism, Econometrica , 76, pp. 31-69.

Cropanzano, R. and M. L Ambrose (2001), Procedural and Distributive Justice Are More Similar Than You Think: A Monistic Perspective and a Research Agenda, Greenberg, J., Cropanzano, R. (eds.), Advances in Organizational Justice, Stanford University Press, 2001.

Dana, J., Weber, R. A., Kuang, J. X. (2007), Exploiting Moral Wiggle Room: Experiments Demonstrating an Illusory Preference for Fairness, Journal of Economic Theory, 33, pp. 67-80.

De Cremer, D., Tyler, T.R., Ouden, N. den (2005), Managing Cooperation Via Procedural Fairness: The Mediating Influence of Self-Other merging, Journal of Economic Psychology, 26, pp. 393-406.

de Tocqueville, A. (1868), De la Dmocratie en Amrique, Prémier Tôme, Quinziéme Edition, Paris: Paguerre.

Dufwenberg, M., Kirchsteiger, G. (2004), A Theory of Sequential Reciprocity Games and Economic Behavior, 47, pp. 268-98.

Engelmann, D., Stobel, M. (2004), Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments, 94, pp. 857-869.

Falk, A., Fehr, E., Fischbacher, U. (2003), On the Nature of Fair Behavior, Economic Inquiry, 41, 20-26.

Falk, A., Fischbacher, U. (2006), A Theory of Reciprocity, Games and Economic Behavior, 54, pp. 293-315.

Fehr, E., Schmidt, G. (1999), A Theory of Fairness, Competition and Cooperation, Quarterly Journal of Economics, 114, pp. 817-868.

Fehr, E., Herz, H., Wilkening, T. (2013), The Lure of Authority: Motivation and Incentive Effects of Power, American Economic Review, 103(4), 1325-59.

Fischbacher, U. (2007), z-Tree: Zurich Toolbox for Ready-made Economic Experiments, Experimental Economics, 10, pp. 171-178.

Frey, B., Stutzer, A. (2005), Beyond Outcomes: Measuring Procedural Utility, Oxford Economic Papers 57, pp. 90-111.

Gehrig, T., Güth, W., Levati, V., Levinsky, R., Ockenfels, A., Uske, T., Weiland, T. (2007), Buying a Pig in a Poke: An Experimental Study of Unconditional Veto Power, Journal of Economic Psychology, 28, 692-703.

Greiner, B. (2004), An Online Recruitment System for Economic Experiments, Kremer, K., Macho, V. (Eds.) (2004), Forschung und wissenschaftliches Rechnen 2003, GWDG Bericht 63, Ges. f. Wiss. Datenverarbeitung, Göttingen.

Grimalda, G., Kar, A., Proto, E. (2012) Everyone Wants a Chance: Initial Positions and Fairness in Ultimatum Games, The Warwick Economics Research Paper Series, # 989.

Güth, W., Schmittberger, R., Schwarz, B. (1982), An Experimental Analysis of Ultimatum Bargaining, Journal of Economic Behavior and Organization, 1982, pp. 367-388.

Güth, W., Tietz, R. (1986), Auctioning Ultimatum Bargaining Positions – How to Act if Rational Decisions are Unacceptable –, in Scholz, R.W. (1986) (Ed.), Current Issues in West German Decision Research, P. Lang Publisher, Frankfurt, Bern, New York, 1986.

Gul, F., Pesendorfer, W. (2005), The Case for Mindless Economics, Caplin, A., Schotter, A. (eds.), The Foundations of Positive and Normative Economics, Oxford University Press, 2008.

Heckman, J. (1979), Sample Selection Bias as a Specification Error, Econometrica, 47(1), pp. 153-161.

Hegel, F. (1821), Grundlinien der Philosophie des Rechts, Berlin: Nicolaische Buchhandlung.

Ishida, C. (2006), How do Scores of the DIT and the MJT differ? A Critical Assessment of the Use of Alternative Moral Development Scales in Studies of Business Ethics, Journal of Business Ethics, 67, pp. 63-74.

Kalai, E., Stanford, W. (1988), Finite rationality and interpersonal complexity in repeated games. Econometrica 56(2), pp. 397-410.

Karni, E., Safra, Z. (2002), Individual Sense of Justice: A Utility Representation, Econometrica, 70, pp. 263-284.

Karni, E., Salmon, T., Sopher, B. (2008), Individual Sense of Fairness: An Experimental Study, Experimental Economics, 11, pp. 174-189.

Kircher, P.A.T., Ludwig, S., Sandroni, A. (2009), Fairness: A Critique to the Utilitarian Approach, SFB/TR 15 Governance and the Efficiency of Economic Systems, Free University of Berlin, Humboldt University of Berlin, University of Bonn, University of Mannheim, University of Munich.

Kohlberg, L. (1969), Stage and Sequence: the Cognitive-Developmental Approach to Socialization, Goslin, D.A. (ed.), Handbook of Socialization and Endash, Theory and research. Chicago: McNally.

Kohlberg, L. (1984), The Psychology of Moral Development. San Francisco: Harper & Row.

Konow, J. (2003), Which is the Fairest One of All? A Positive Analysis of Justice Theories, Journal of Economic Literature, 41, pp. 1188-1239.

Krawczyk, M. W. (2011), A Model of Procedural and Distributive Fairness, Theory and Decision 70, pp. 111-128.

Krawczyk, M. W., Le Lec, F. (2010), Give Me a Chance: An Experiment on Social Decision Making, Experimental Economics, 13, pp. 500-511.

Leventhal, G.S. (1976), The Distribution of Rewards and Resources in Groups and Organizations, L. Berkowitz and E. Walster (eds.), Advances in Experimental Social Psychology Vol. 9, NY: Academic Press.

Lind, G. (1978), Wie misst man moralisches Urteil? Probleme und alternative Möglichkeiten der Messung eines komplexen Konstrukts, Portele, G. (ed.), Sozialisation und Moral, pp. 171-201, Weinheim: Beltz.

Lind, G. (2000), The Importance of Role-Taking Opportunities for Self-Sustaining Moral Development, Journal of Research in Education, 10, pp. 9-15.

Lind, G. (2002), Ist Moral lehrbar? Ergebnisse der Modernen Moralpsychologischen Forschung, 2nd Edition, Berlin: Logos-Verlag.

Lind, E. A. (2001), Fairness Heuristic Theory: Justice Judgements as Pivotal Cognitions in Organizational Relations, Greenberg, J., Cropanzano, R. (eds.), Advances in Organizational Justice, Stanford University Press, 2001.

Lind, G. (2008), The Meaning and Measurement of Moral Judgement Competence Revisited - A Dual-Aspect Model, Fasko, D., W. Willis (eds.), Contemporary Philosophical and Psychological Perspectives on Moral Development and Education. Cresskill, NJ: Hampton Press, pp. 185 - 220.

Miettinen, T. (2013), Promises and Conventions: An Approach to Pre-play Agreements, Games and Economic Behavior, 80, pp. 63-84.

Neri, C., Rommeswinkel, H. (2014), Freedom, Power, and non-Interference: An Experiment on Decision Rights, mimeo, http://www.eea-esem.com/files/papers/EEA-ESEM/2014/387/Neri_Rommeswinkel.pdf.

Osbourne, M., and Rubinstein, A. (1994), A Course in Game Theory, MIT press.

Piaget, J. (1948), The Moral Judgement of the Child, Glencoe, Illinois: Free Press.

Price, P.C. (1998), Effects of a Relative-Frequency Elicitation Question on Likelihood Judgment Accuracy: The Case of External Correspondence, Organizational Behavior and Human Decision Processes 76(3), pp. 277-297.

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

Rabin, M. (2002), A Perspective on Psychology and Economics, European Economic Review 46, pp. 657-685.

Rawls, J. (1958), Justice as Fairness, The Philosophical Review, 67(2), pp. 164-194.

Rawls, J. (1971), A Theory of Justice, Cambridge, Harvard University Press.

Rosenbaum, P. R. and Rubin, D. B. (1983), The Central Role of The Propensity Score in Observational Studies for Causal Effects, Biometrika, 70, pp. 41-55.

Rubinstein, A. (1986), Finite automata play the repeated prisoner's dilemma, Journal of Economic Theory, 39(1), pp. 83-96.

Schlag. K., Tremewan, J. (2012), Simple Belief Elicitation, University of Vienna Department of Economics Working Paper 1204.

Schlag. K., Tremewan, J., von der Weele, J. (2014), A Penny For Your Thoughts: A Survey of Methods Eliciting Beliefs, Manuscript, University of Vienna.

Sebald, A. (2010), Attribution and Reciprocity, Games and Economic Behavior, 68, pp. 339-352.

Selten, R. (1967), Die Strategiemethode zur Erforschung des eingeschränkt rationalen Verhaltens im Rahmen eines Oligopolexperiments, Sauermann, H. (ed.), Beiträge zur experimentellen Wirtschaftsforschung, J.C.B. Mohr (Paul Siebeck), Tübingen, pp. 136-168.

Shor, M. (2009), Procedural Justice in Simple Bargaining Games, Manuscript, Vanderbilt University.

Sondak, H. and Tyler, T. R. (2007), How does Procedural Justice Shape the Desirability of Markets? Journal of Economic Psychology, 28, pp. 79-92.

Sugden, R. (1998), The Metric of Opportunity, Economics and Philosophy, 14, pp. 307-337.

Thibaut, J. and Walker L. (1975), Procedural Justice: A Psychological Analysis. Hillsdale, NJ: Erlbaum.

Trautmann, S. (2009), A tractable Model of Process Fairness under Risk, Journal of Economic Psychology 30, pp. 803-813.

Tyler, L. (2000), Social Justice: Outcome and Procedure, International Journal of Psychology, 35, pp. 117-125.

Tyler T. (2004), Procedural Justice, A. Sarat (ed.) The Blackwell Companion to Law and Society, Malden, MA: Blackwell. pp. 435-452

Tyran J.-R., Sausgruber, R. (2014), Discriminatory Taxes are Unpopular – Even When They Are Efficient and Distributionally Fair, Journal of Economic Behavior and Organization, http://dx.doi.org/10.1016/j.jebo.2013.12.022.

Wasel W. (1994), Simulation moralischer Urteilsfähigkeit. Moralentwicklung: eine kognitiv-strukturelle Veränderung oder ein affektives Phänomen?, Universität Konstanz, Diplomarbeit im Fachbereich Psychologie, http://www.uni-konstanz.de/ag-moral/mut/mjt-references.htm#wasel_1994.

Weber, M. (1921), Wirtschaft und Gesellschaft: Grundriß der verstehenden Soziologie, J.C.B. Mohr (Paul Siebeck), Tübingen 1921; Roth, G. Wittich, C. (Eds., Transl.) (1968), Economy and Society, New York: Bedminster Press.

A Appendices

A.1 Instructions[43]

Instructions

Welcome and thank you for participating in this experiment. For showing up on time you receive €2.50. Please read the following instructions carefully. Instructions are identical for all participants. Communication with other participants must cease from now on. Please switch off your mobile phone.

If you have any questions, raise your hand – we will answer them individually at your seat.
During the experiment all amounts will be stated in ECU (Experimental Currency Units). The sum of your payoffs from all rounds will be disbursed to you in cash at the end of the experiment (exchange rate: 1 ECU=0.03 €). Your initial endowment is 20 ECU.

Information regarding the experiment

Participants take on different roles **A** and **B**. You do not know your role in the beginning and will at first make decisions for role A as well as for role B. You will then be randomly assigned one role and will be informed accordingly. From then on, roles remain the same throughout the experiment.

You will be randomly matched with other anonymous participants. Your decisions affect your own payoff and the payoffs of those participants with whom you interact.

In the experiment, you encounter two situations. These situations are characterized as follows:

Situation 1. There are **200 ECU**. Participant **A** chooses between two alternatives **X** and **Y** to allocate these 200 ECU between herself and participant **B**.

X: She allocates **100 ECU** to herself and **100 ECU** to participant **B**.
Y: She allocates **20 ECU** to herself and **180 ECU** to participant **B**.

Participant B does not learn A's choice. B chooses between **U** and **V**:

U: **B agrees** with the allocation **unknown to her**. If so, the allocation corresponds to participants' payoffs in ECU.

V: **B does not agree** with the allocation **unknown to her**. If so, both participants obtain a payoff of **0 ECU**.

[43]Instructions of the experiment were written in German. The following chapter reproduces a translation into English for experimental sessions which introduced the Ultimatum and the Yes-no game. Emphases in bold or italic font are taken from the original text, **TEXT IN CAPITAL LETTERS WAS NOT PART OF THE ORIGINAL INSTRUCTIONS**. Instructions for other treatments are available from the authors.

Situation 2. There are **200 ECU**. Participant **A** chooses between options **X** and **Y** to allocate these 200 ECU between herself and participant **B**.

X: She allocates **100 ECU** to herself and **100 ECU** to participant **B**.
Y: She allocates **20 ECU** to herself and **180 ECU** to participant **B**.

Participant B learns A's choice and chooses between **U** and **V**.

U: **B agrees** with the allocation **known to her**. If so, the allocation corresponds to participants' payoffs in ECU.

V: **B does not agree** with the allocation **known to her**. If so, both participants obtain a payoff of **0 ECU**.

All participants now make their decisions for both roles and for both situations. You state for role A which option (**X** or **Y**) you would choose in situation 1 and situation 2, respectively. For role B, you decide for every situation between **U** and **V**. Both situations are initialized to occur with equal probability 0.50 (50%). The decisions made for the situation which is drawn become payoff relevant. Payoffs are calculated as described above.

Please be patient until the experiment starts. If you have any questions, raise your hand. Before the experiment starts, please answer the following control questions.

A.2 Control Questions

Control Questions[44]

1. Assume that participants choose as follows:

   participant A:

situation 1	situation 2
X	X

   participant B:

situation 1	situation 2	
	if X	if Y
agrees	agrees	disagrees

   This means that in situation 1 and in situation 2, participant A chooses X. Participant B agrees in situation 1. In situation 2, she agrees if A chooses X, and she does not agree if A chooses Y. If situation 1 is chosen randomly, what is (in ECU)

   (a) participant A's payoff?

   (b) participant B's payoff?

   If situation 2 is chosen randomly, what is (in ECU)

[44] CONTROL QUESTIONS ABOUT THE ACTIONS AND SITUATIONS IN PHASE 1.

(a) participant A's payoff?

(b) participant B's payoff?

2. Assume that A and B still choose as described in 1., with the exception that in situation 2, A now chooses Y.

   (a) What is participant B's payoff in situation 2?

Please press 'OK'.

3. What is the difference between situation 1 and 2? Please choose 'right' or 'wrong'.

   (a) In situation 2, B has two courses of action whereas in situation 1, she has one.

   (b) Both in situation 1 and in situation 2, B knows which distribution of payoffs A has chosen.

   (c) In situation 2, B can actually react to A's choice whereas in situation 1, she can merely make a decision.

Please press 'OK'.

A.3 Instructions – Bidding Phase

INSTRUCTIONS – BIDDING PHASE

Now, one of either participant may influence which situation is drawn. This participant is determined by casting lots between participant A and participant B. Thereby, A and B have an equal chance to be drawn. If drawn by chance, a participant can pay the amount of 5 ECU to make the situation she prefers (if any) more likely to occur. If she does not pay, both situations occur as they have been initialized with 50 % probability. At the end of the experiment, one situation will be drawn. The decisions made for this situation become payoff-relevant.

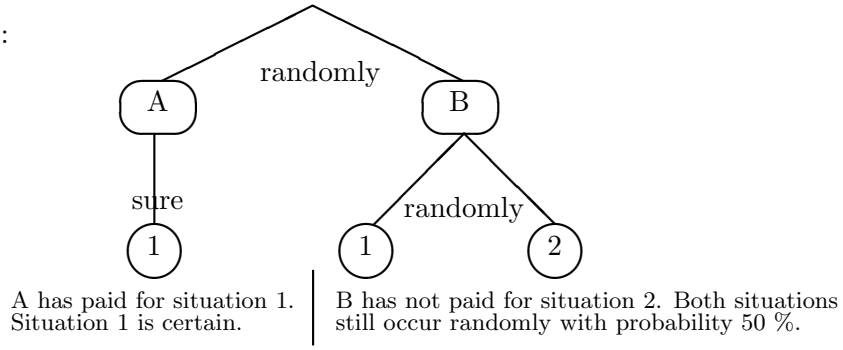
Payoffs for each situation are calculated as described in the instructions. If you may influence the draw of the situations and choose to do so, the cost of influencing the draw of the situations will be deducted from this payoff.

A.4 Control Questions – Bidding Phase

Control Questions[45]

Assume that A preferred situation 1 and paid 5 ECU for this situation. B preferred situation 2 but did not pay for this situation. Chance has not yet decided which participant's choice will actually be implemented. How likely is it that situation 1 occurs?

Some graphical help:



Please choose 'right' or 'wrong':

1. Situation 1 is certain. right/wrong.
2. Situation 1 is more likely than situation 2 (but not certain). right/wrong.
3. Situation 1 is as likely as situation 2. right/wrong.
4. Situation 1 is less likely than situation 2 (but not impossible). right/wrong.
5. Situation 1 is impossible. right/wrong.

Please press 'OK'. *(SUBJECTS ALSO HAD THE POSSIBILITY TO GO BACK TO THE PREVIOUS SCREEN WHICH SHOWED THE INSTRUCTIONS FOR THE BIDDING PHASE – SEE ABOVE.)*

[45] ABOUT THE INSTRUCTIONS FOR PHASE 2, I.E. THE BIDDING MECHANISM.

A.5 An Excerpt of the Moral Judgement Test by Georg Lind (1976, 2008)

Workers

Recently a company fired some people for unknown reasons. Some workers think that their bosses are listening in on their private conversations through cameras and microphones in the building and using the information against them. The bosses say they are not listening in	The workers cannot legally do anything until they can prove that their bosses are listening in on their conversations. Two workers then break into the main office and take the tapes that prove their bosses were listening in.

Would you agree or disagree with the workers' action ...

I strongly disagree — I strongly agree

-3	-2	-1	0	1	2	3

How acceptable do you find the following arguments *in favor* of the two workers' actions?
Suppose someone argued they were *right* for breaking in ...

...because most of the workers would approve of their actions and many would be happy about it.

I strongly reject — I strongly accept

-4	-3	-2	-1	0	1	2	3	4

. . .

...because the company had done something wrong first by listening in, the two workers were right in breaking into the main office.

I strongly reject — I strongly accept

-4	-3	-2	-1	0	1	2	3	4

. . .

How acceptable do you find the following arguments *against* the two workers' actions?
Suppose someone argued they were *wrong* for breaking in ...

. . .

...because a person doesn't steal if he wants to be considered decent and honest.

I strongly reject — I strongly accept

-4	-3	-2	-1	0	1	2	3	4

. . .

...because the two workers should have used all the legal options available to them without breaking the law.

I strongly reject — I strongly accept

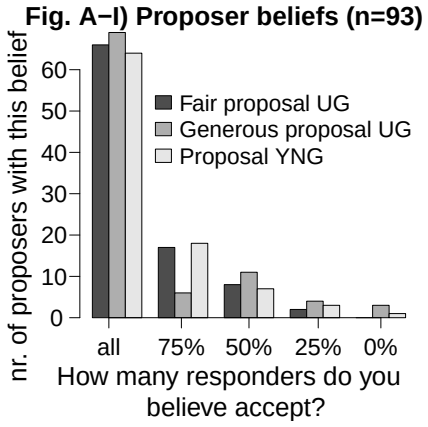
-4	-3	-2	-1	0	1	2	3	4

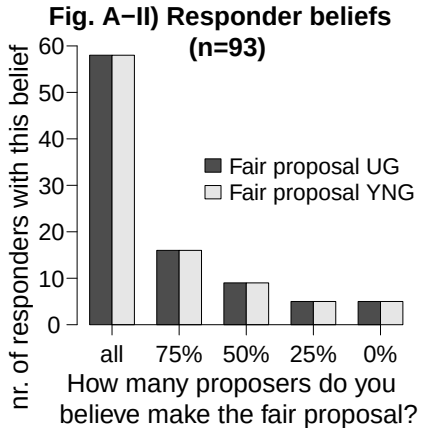
. . .

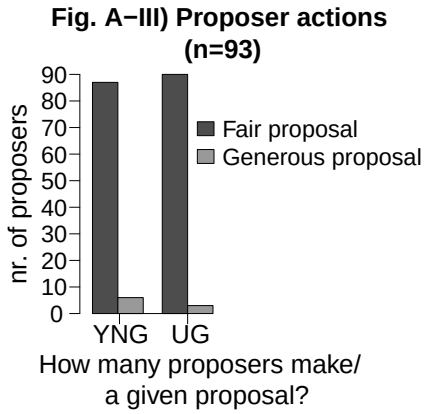
NOTE: This excerpt of the moral judgement test MJT is reprinted with kind permission by Georg Lind. It does not faithfully reproduce the formatting of the original test. For ease of readability, the original test numbers each item, and the alignment slightly differs from this excerpt. The dots represent items which have been left out. The full test cannot be published due to copyright issues.

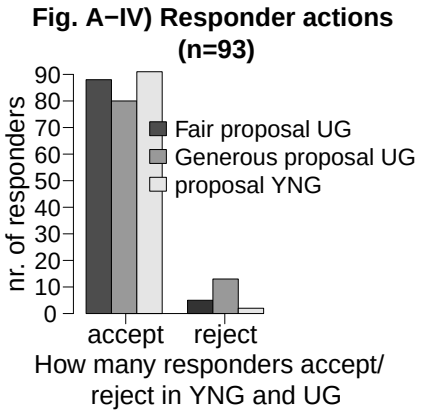
B Overall behavior and beliefs across protocols

B.1 Ultimatum vs. Yes-No Game

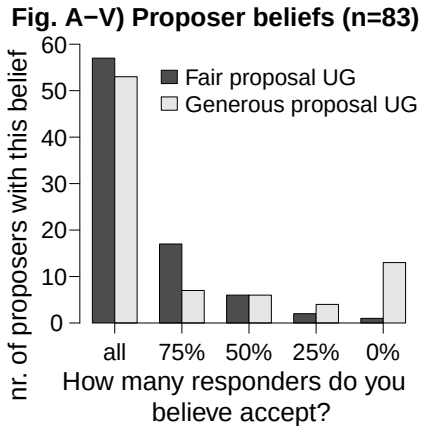


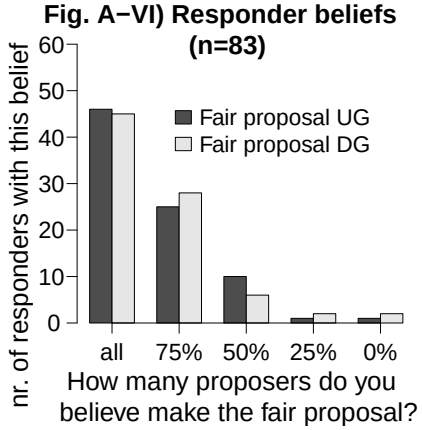


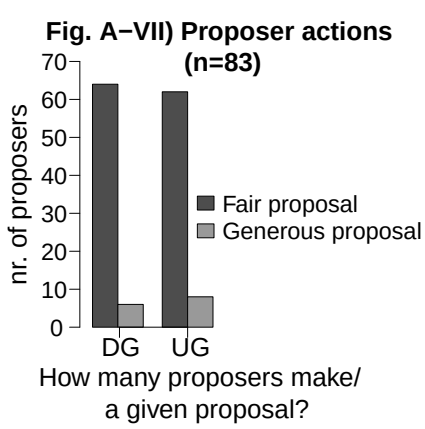


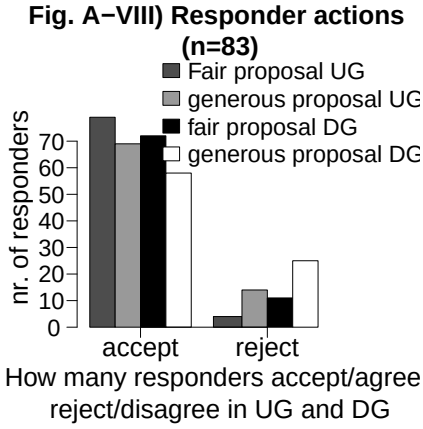


B.2 Dictator vs. Ultimatum Game









C Lawrence Kohlberg's six 'classes' or 'ways' of argumentation.

Table A1: Six ways of moral argumentation (summary by Ishida 2006, examples from the authors).

argumentation	Classes of motivation for moral behavior	I prefer...
preconventional way	***Kohlberg 1.*** Orientation to punishment and obedience, physical and material power. Rules are obeyed to avoid punishment. ***Kohlberg 2.*** Naïve hedonistic orientation. The individual conforms to obtain rewards.	...the yes-no game because therein, I will not be punished for not being generous./ ...the ultimatum game: because the responder can and will reward me for being generous by accepting the proposal.
conventional way	***Kohlberg 3.*** "Good boy/girl" orientation to win approval and maintain expectations of one's immediate group. The individual conforms to avoid disapproval. One earns approval by being "nice".	...the ultimatum game because therein, I can signal my generous intentions to the responder who will reciprocate by accepting/
	Kohlberg 4. Orientation to authority, law, and duty, to maintain a fixed order. Right behavior consists of doing one's duty and abiding by the social order.	...because the responder expects me to be generous, and in the ultimatum game, I can show the responder I do not want to disappoint her expectations and let her down...
postconventional way	***Kohlberg 5.*** Social contract orientation. Duties are defined in terms of the social contract and the respect of others' rights. Emphasis is upon equality and mutual obligation within a democratic order.	...the yes-no game: it is more democratic since it grants both parties equality in decision and information rights/...the ultimatum game: it proceeds more transparently and the social contract can only be backed by transparent institutions/
	Kohlberg 6. The morality of individual principles of conscience, such as the respect for the individual will, freedom of choice etc. Rightness of acts is determined by conscience in accord with comprehensive, universal and consistent ethical principles.	...the ultimatum game: as proposer, I respect the responder's will and she has more opportunity to express this will in the ultimatum game

D Purely procedural concerns

D.1 Inequality in information: Formalization

As before, we use the terminology of Osbourne and Rubinstein (1994) if not otherwise stated. Let Γ be a two-player extensive form game where each player moves at most once. Let $s_i \in S_i$ be a strategy of player i in her strategy set in that game. A terminal history of the game in the set of terminal histories is denoted by $z \in Z$.

If we wish to model players who care about the interpersonal dimension in the distribution (or put differently, the precision) of information, we first need a means to express the amount of information each player has. There are two sources of information for a player: first, information about events exogenous to the game (e.g. information about nature's move) that each player has. Second, the information which each player learns about her opponent's actions. We assume here that each player can perfectly control and learn her own actions, and also assume perfect recall. Information from both sources determines how well a player can predict which terminal node or history of a game will be reached. If both players can transparently observe all actions and gain all relevant information about exogenous events and all actions at each stage of a procedure, then each player knows the terminal history for sure and coincidentally, there is also equity of information (there is also equity of information if players ignore the terminal history of the game to the same extent). If one of the players knows all relevant aspects and controls all decisions determining the allocation of material benefits in the game and this takes place without any transparency or possibilities for the opponent to monitor those actions, then there is severe asymmetry of information about the terminal histories of the game. Hence, we express the amount of information for each player by the fragmentation of her information partition about the terminal histories of the game. These information partitions have, to date, not directly entered the utility function, and thus not been modelled as directly relevant for indvidual preferences.

Let us denote player $j's$ partition of information over the terminal nodes with $\mathcal{I}_j^z$. This is what j knows about terminal nodes given j's own information, what j learns about $i's$ actions, and the control j has over her own actions when she is active. These partitions for players 1 and 2, respectively, will in a natural way be perfectly determined by the player nodes, information partitions, and action sets for each player.

As examples, consider the ultimatum game and the yes-no game. In both games, both players fully control their own actions: the proposer fully controls her proposal, the responder fully controls her acceptance/rejection decision. Yet, the two games differ regarding how much the responder knows about the proposal. In the ultimatum game, the responder learns the proposal made by the proposer. Since in addition, the responder also controls her own decision, she knows which terminal node will be reached. Therefore, the four terminal nodes of the ultimatum game are partitioned into singleton sets for the responder. The proposer in turn fully controls her own action – the proposal she makes/made. She does, however, not know how the responder reacts to each of her two potential proposals. Thus, the proposer's information partition over the terminal nodes consists of two non-singleton sets each containing two terminal nodes: the first set contains the responder's acceptance and rejection of the

fair proposal; the second set containing the acceptance and rejection of the generous proposal. In summary, the cardinality of the information partitions over the terminal nodes of the ultimatum game are 2 for the proposer, and 4 for the reponder, respectively. In the yes-no game, the responder does not learn the proposal. She fully controls her acceptance/rejection. Thus, her partition over the terminal histories of the game contains two sets, i.e. has *cardinality two*: one set with the two possible terminal histories where the responder has accepted, another set with the two terminal histories where she has rejected. The proposer's information partition is identical in the yes-no and the ultimatum games, since she controls the proposal, but does not know how the responder will react. The information partition has therefore *cardinality two* as well.

Using these measures for how much information each player has, we can now express a player's aversion to information asymmetries. If player i cares about purely procedural fairness and the equality of access to information in particular, her preferences could be characterized by the utility function

$$u_i(s_i, s_j; b_i, b_j) - \beta_i max\{\#\mathcal{I}_i^z - \#\mathcal{I}_j^z, 0\} - \alpha_i max\{\#\mathcal{I}_j^z - \#\mathcal{I}_i^z, 0\}$$

where $u_i(s_i, s_j; b_i, b_j)$ captures the social welfare function dependent on the outcome s_i, s_j (as in inequity aversion models; Fehr-Schmidt, 1999; Bolton and Ockenfels, 2000, for instance) and possibly on players' belief systems b_i, b_j (as in psychological games; Battigalli and Dufwenberg, 2009). The procedural fairness notion of inequity aversion in access to information is modelled as $-\beta_i max\{\#\mathcal{I}_i^z - \#\mathcal{I}_j^z, 0\} - \alpha_i max\{\#\mathcal{I}_j^z - \#\mathcal{I}_i^z, 0\}$ where the first term captures the aversion for advantageous inequality in access to information and the latter term the aversion for disadvantageous inequality in access to information. Notice that the cardinality of a set B, $\#B$, denotes the number of elements in that set. This is the simplest specification with piecewise linear utility in information asymmetries. As an analogy with Fehr-Schmidt (1999), it is natural to assume that $\alpha_i \geq \beta_i$ so that players are assumed to be more aversive to disadvantageous inequality than to advantageous inequality. Thus a proposer and a responder with identical procedural preferences facing a choice between the same two procedures may each prefer a different procedure just, because of their role, the inequality in access to information in a given procedure is advantageous for one of the players and disadvantageous for the other (see tables 3 and 4 in section 6). Such a proposer would have a payoff $u_1^{UG}(s_1, s_2; b_1, b_2) - \alpha_1 max\{\#\mathcal{I}_2^{z,UG} - \#\mathcal{I}_1^{z,UG}, 0\} = u_1^{UG}(s_1, s_2; b_1, b_2) - \alpha_1 max\{4-2, 0\}$ in the ultimatum game, and a payoff $u_1^{YNG}(s_1, s_2; b_1, b_2) - \alpha_1 max\{\#\mathcal{I}_2^{z,YNG} - \#\mathcal{I}_1^{z,YNG}, 0\} = u_1^{YNG}(s_1, s_2; b_1, b_2) - \alpha_1 max\{2-2, 0\}$ in the yes-no game. Thus the proposer with purely procedural concerns of equality of information would strongly prefer the yes-no game if the terms $u_1^{UG}(s_1, s_2; b_1, b_2)$ and $u_1^{YNG}(s_1, s_2; b_1, b_2)$ are equal (which requires analogous actions and beliefs in the two procedures, see table 2, section 3). The responder with purely procedural concerns of equality of information would also prefer the yes-no game but her preference would be weaker since $\alpha \geq \beta$. This is in line with the observed revealed preference patterns over the two procedures (see table 4, section 6).

D.2 Procedural transparency

Hegel (1821, §215) argues that people should have an equal claim to jurisprudence which can only be the case if the law is transparent, and in particular, if all decisions pertaining to judicial processes are common knowledge to all parties at all points in time. Rawls (1958) argues that transparency – along with simplicity, and equal freedom of choice – define fairness which in turn promotes justice. Transparency of institutions, does, therefore, also have ethical content. Moreover, as a necessary feature of those institutions backing the social contract, it could be motivated by the same ethical ideal from which preferences about the equality of rights should spring, that is, Kohlberg class 5, see table 5. There are two games which proceed transparently in our setting: the dictator, and the ultimatum game. Whenever a party is called upon to choose, she knows all decisions which have previously been made. Note that subjects who choose between the yes-no and the ultimatum game can only opt for transparency (i.e. the ultimatum game) at the cost of introducing unequal information and decision rights.

D.3 Procedural simplicity

We express the simplicity of a procedure by the number of eventualities a player needs to reason about, see already (de Tocqueville 1868) for some aspects, and the desirability of this property[46]. This number of eventualities depends on two elements: the number of the opponent's choices, and the number of the player's own choices. For each opponent choice, the player must determine what her own preferred reaction to this choice is, and whether given this reaction, the opponent choice was in the opponent's interest given some preference the opponent might hold. The higher this number of eventualities, the more cognitive effort is required, and the more cognitive resources are bound. Players could prefer procedures where the number of strategic eventualities she needs to consider, is small(er). In the yes-no game, each player has to think about the two moves of her own, and the two moves of the other player. Therefore, each player in a yes-no game has to think about altogether only four possible combinations of moves (which coincides with the cardinality of a player's set of 'pure strategies')[47]. In the ultimatum game, each player has to think about the proposer's two moves, and the responder's two moves given each proposal. Altogether, each player needs to think about six possible combinations of moves. In terms of procedural simplicity, the yes-no game is therefore *simpler* than the ultimatum game. Since the yes-no game also distributed rights equally while the ultimatum game did not, a natural way to disentangle these motivations is to look whether a player's preference for the yes-no game correlates with her moral judgement (motive: distribution of rights), or not (motive: simplicity). Looking at this paper's specific dictator game, proposer and responder also have to think about six eventualities each: the proposer needs to understand that whatever she proposes, whether the responder agrees or disagrees with each proposal, does not change the final

[46]The complexity of strategies has also been described game-theoretically by e.g. Rubinstein (1986) or Kalai and Stanford (1988)

[47]We do not explicitly consider mixed strategies. But note that the pure strategies are the limiting case for each mixing strategy, and therefore, two different sets of distinct pure strategies – whatever they are – always spawn the exact same number of mixed strategies on a continuous scale.

allocation. The responder needs to understand the same.

D.4 Procedural efficiency

In our setting, the proposer can only make a fair, and a generous proposal. Hence, she cannot bias distributive fairness in a self-serving way. The veto right in our mini-ultimatum game thus does not protect the responder from a proposer's self-serving distributive unfairness: the veto is merely an inefficiency-inducing option. Responders and proposers could intrinsically value procedures which preclude conflict, even if they know for sure they agree, and that conflict is a purely hypothetical scenario. In our setting, the only game which meets the criterion of purely procedural efficiency, is the dictator game.

Table A2: Yes-No Game vs Ultimatum Game: Distribution Of Rights across proposer and responder, Simplicity, Transparency, and Efficiency of each game.

	role	yes-no game	ultimatum game
decision rights (nr. of effective pure strategies)	proposer	2	2
	responder	2	4
	distribution of rights	$\{2,2\}$	$\{2,4\}$
information rights (cardinality of information partition over terminal nodes)	proposer	2	2
	responder	2	4
	distribution of rights	$\{2,2\}$	$\{2,4\}$
simplicity (sum of own and opponent's moves a party has to reason about)	proposer responder	4	6
transparency: game has perfect information	proposer responder	no	yes
efficient regulation of conflicts?		no	no

Table A3: Dictator Game vs Ultimatum Game: Distribution Of Rights across proposer and responder, Simplicity, Transparency, and Efficiency of each game.

	role	dictator game	ultimatum game
decision rights (nr. of effective pure strategies)	proposer	2	2
	responder	1	2
	distribution of rights	$\{2,1\}$	$\{2,4\}$
information rights (cardinality of information partition over terminal nodes)	proposer	4	2
	responder	4	4
	distribution of rights	$\{4,4\}$	$\{2,4\}$
simplicity (sum of own and opponent's move a party has to reason about)	proposer responder	6	6
transparency: game has perfect information	proposer responder	yes	yes
efficient regulation of conflicts?		yes	no

E Predictions of the sequential reciprocity equilibrium, Dufwenberg and Kirchsteiger (2004)

Proposition (YNG). *There is a unique equilibrium. The proposer (all types) proposes F. A responder with sensitivity to reciprocity $Y_R \leq 1/40$ accepts with probability one, a responder with $Y_P > 1/40$ accepts with probability $q = \frac{1}{40Y_R}$.*

Proof. The responder has a single efficient strategy (see Dufwenberg and Kirchsteiger, 2004, pp. 276): to accept with probability one. Therefore, the responder R is commonly known to be unkind towards the proposer P. The responder's kindness towards the proposer is captured by variable κ_{RP} where kindness is associated with a positive value and unkindness associated with negative value. By the above argument, $\kappa_{RP} \leq 0$.

Given acceptance rate q, the proposer's pecuniary payoff for proposing F is $100q$ and that for proposing G is $20q$. The responder's respective payoffs are $100q$ and $180q$. The proposer proposes F if the payoff for doing so (on the left-hand side of the following inequality) is greater than the payoff of proposing G (on the right-hand side)

$$100q + Y_P\kappa_{RP}(100q - \frac{100q+180q}{2}) > 20q + Y_P\kappa_{RP}(180q - \frac{100q+180q}{2})$$

where the parameter Y_P is the proposer's sensitivity to reciprocity, $(100q - \frac{100q+180q}{2})$ and $(180q - \frac{100q+180q}{2})$ measure the proposer's kindness κ_{PR} of proposing F and G, respectively. Since κ_{RP} is non-positive, the responder maximizes her payoff by proposing F.

The responder accepts if the payoff of accepting (the left-hand side of the following inequality) is greater than that of rejecting (on the right hand side)

$$100 + Y_R \times 0 \times \kappa_{PR} > 0 + Y_R \times (-100) \times \kappa_{PR}$$

where $\kappa_{PR} = \frac{100q-180q}{2} < 0$. The inequality simplifies to$Y_R < \frac{1}{40q}$. If to the contrary $Y_R > \frac{1}{40q}$, then the responder rejects the fair proposal. Notice that in equilibrium, the proposer must have correct beliefs about the rejection rate. Thus, in equilibrium the responder never rejects with probability one. The responder with sensitivity to reciprocity $Y_R \leq 1/40$ accepts with certainty and a responder of specific sensitivity $Y_R = \frac{1}{40q}$ is indifferent and accepts with probability $q = \frac{1}{40Y_R}$. *QED.*

Proposition (UG). *Under the restriction $q_F = q_G$, there is a unique equilibrium where $q_F = q_G = 1$. The proposer (all types) proposes F. A responder with sensitivity to reciprocity $Y_R \leq 1/40$ and accepts with probability one. (The proposer must expect $Y_R \leq 1/40$ with probability one).*

Proof. As in the yes-no game, the responder can only be neutral or unkind, $\kappa_{RP} \leq 0$. Given the acceptance rates q_F and q_G of the fair and the generous proposal respectively, the proposer's pecuniary payoff for proposing F is $100q_F$ and that for proposing G $20q_G$. The responder respective payoffs are $100q_F$ and $180q_G$. The proposer proposes F if $100q_F + Y_P\kappa_{RP}(100q_F - \frac{100q_F+180q_G}{2}) >$

$20q_G + Y_P\kappa_{RP}(180q_G - \frac{100q_F+180q_G}{2})$, i.e. if

$$100q_F - 20q_G > Y_P\kappa_{RP}[180q_G - 100q_F].$$

Three cases: (1) $q_G < 5/9q_F$. In this case, the proposer prefers F if

$$Y_P < \frac{100q_F - 20q_G}{\kappa_{RP}(180q_G - 100q_F)}.$$

(2) $5q_F \geq q_G \geq 5/9q_F$. (this includes the case $q_F = q_G$). In this case, the proposers of all sensitivities Y_P prefer F. (3) $5q_F < q_G$.In this case the proposer prefers F if $Y_P > \frac{100q_F-20q_G}{\kappa_{RP}(180q_G-100q_F)}$.

We are interested in predictions under the restriction that the responder is expected to accept both proposals with equal probability, $q_F = q_G$ (this is something we control for by eliciting beliefs). In this case the proposer always proposes F. The responder who expects that the fair proposal is proposed accepts if $Y_R < \frac{1}{40q_F}$. By the same argument as above, the responder accepts with certainty if $Y_R < \frac{1}{40q_F}$, i.e. in equilibrium where beliefs are correct $Y_R < \frac{1}{40}$. There is no pure strategy equilibrium where the responder rejects with certainty. Yet, given a commonly known sensitivity type Y_R, there is a mixed strategy equilibrium where the type $Y_R = \frac{1}{40q_F}$ is indifferent and accepts with probability $q_F = \frac{1}{40Y_R}$).

Let us finally verify that it is optimal to accept G with probability $q_G = q_F$. Acceptance is preferred if

$$180 + Y_R \times 0 \times \kappa_{PR} > 0 + Y_R \times (-20) \times \kappa_{PR}$$

where $\kappa_{PR} = \frac{180q-100q}{2} > 0$ and thus acceptance is always preferred. The unique equilibrium under our restriction $q_F = q_G = 1$ where responder is of type $Y_R \leq 1/40$. *QED.*

Proposition (Procedural indifference). *If $q_F = q_G = 1$, each player is indifferent between whether UG or YNG is used/played.*

Proof. If $q_F = q_G = 1$, the proposer proposes F and the responder accepts with certainty. Thus. the responder's equilibrium payoff equals $100 + Y_R \times \kappa_{RP} \times \kappa_{PR}$ where both in the YNG and in the UG, $\kappa_{RP} = 0$ (the responder is neither kind or unkind). Thus the expected payoffs are equal in both games. It is easy to verify that the same argument implies that also the proposer payoffs are equal in the two games.

In the dictator game, the responder cannot influence the payoffs, so he can only be neutral $\kappa_{RP} = 0$. Thus the proposer receives the same payoff in the UG and in the DG, so does the responder. Therefore, there is procedural indifference between the two procedures if $q_F = q_G = 1$. *QED.*

F Non-'EQ'-subjects: strategic incentives and procedural choices

F.1 Yes-No vs Ultimatum Game

Table A4: Yes-No vs Ultimatum game: Strategic incentives, and actual procedural choices for both roles and all clusters from section 7.2

role	cluster nr. (# nr of observations in brackets)	game preference (#nr of observations in brackets)	material advantage *where?*	material advantage *size*	payment[48]
proposer	1 & 2 (#31)	indifference (#10)	ultimatum	11.50	cannot pay
		yes-no (#15)	ultimatum	1.00	9/15
		ultimatum (#6)	ultimatum	14.17	2/6
	3 (#17)	indifference (#6)	ultimatum	40	cannot pay
		yes-no (#10)	yes-no	9	5/10
		ultimatum (#1)	ultimatum	50	1/1
responder	1 (#22)	indifference (#7)	ultimatum	9.29	cannot pay
		yes-no (#6)	ultimatum	98.33	3/6
		ultimatum (#9)	ultimatum	26.67	1/9
	2&3 (#33)	indifference (#12)	yes-no	32.08	cannot pay
		yes-no (#9)	yes-no	22.78	6/9
		ultimatum (#12)	yes-no	33.75	4/12

[48] Reading example: Take the first line of table A4. The first cluster we analyzed in section 7.2 was a merger between cluster 1 with n=22 and cluster 2 with n=9. In the merged cluster, 10 subjects state they are indifferent. These 10 subjects believe they have a material advantage in the ultimatum game (see column 4.1) of an average 11.50 ECU (see column 4.2). Since only subjects who state a positive preference for one game can pay, these 10 subjects cannot pay (see column 5) to influence the draw of the procedures. Take the second line. 15 subjects state to prefer the yes-no game. On average, they believe to have a slight average advantage in the ultimatum game of 1 ECU. 9 out of them actually pay for the yes-no game. Hence, for this group, neither the stated preference, nor the payment decision are in line with their material incentive. Note also that for these subjects, those who pay and those who do not pay *reveal* whatever they state to prefer: both forego payoff but those who pay forego more than those who do not.

Table A5: Dictator vs Ultimatum game: Strategic incentives, and actual procedural choices for efficiency-minded, and non-efficiency minded individuals opting for the dictator game; for both roles and all clusters in section 7.2

role	cluster nr. (# nr of observations in brackets)	game preference (#nr of observations in brackets)	motive[49]	material advantage *where?*	material advantage *size*	payment
proposer	1, $n = \#24$	indifference (#1)	(-)	dictator	80	cannot pay
		dictator (#21)	efficiency (#6)	dictator	44.17	1/6
			other (#15)	dictator	24.67	8/15
		ultimatum (#2)	(-)	dictator	55	1/2
	2, $n = \#24$	indifference (#1)	(-)	dictator	80	cannot pay
		dictator (#19)	efficiency (#7)	ultimatum	9.29	6/7
			other (#12)	dictator	11.67	5/12
		ultimatum (#2)	(-)	dictator	95	2/2
responder	1, #33	indifference (#7)	(-)	dictator	21.43	cannot pay
		dictator (#22)	efficiency (#6)	dictator	10	5/6
			other (#16)	dictator	30.94	9/16
		ultimatum (#4)	(-)	dictator	40	3/4
	2, #13	indifference (#4)	(-)	dictator	10	cannot pay
		dictator (#7)	efficiency (#3)	ultimatum	20	2/3
			other (#4)	ultimatum	20	2/4
		ultimatum (#2)	(-)	dictator	17.50	0/2
	3, #12	indifference (#3)	(-)	dictator	20	cannot pay
		dictator (#4)	efficiency (#2)	ultimatum	5	1/2
			other (#2)	dictator	25	2/2
		ultimatum (#5)	(-)	ultimatum	18	3/5

[49] We only classified whether a subject who had opted for the dictator game, had stated an efficiency reason in the open form section of the post-experimental questionnaire, or not. Subjects who chose the ultimatum game or stated indifference do therefore have no entries in the 'motive' table.