Decision-making Procedures

Aldashev, Gani; Kirchsteiger, Georg; Sebald, Alexander Christopher

DISCUSSION PAPER SERIES

No. 7365

DECISION-MAKING PROCEDURES: A GENERAL THEORY AND ITS FIELD EXPERIMENTAL TEST

Gani Aldashev, Georg Kirchsteiger
and Alexander Sebald

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP7365.asp**

ISSN 0265-8003

DECISION-MAKING PROCEDURES: A GENERAL THEORY AND ITS FIELD EXPERIMENTAL TEST

Gani Aldashev, FUNDP
Georg Kirchsteiger, ECARES, Université Libre de Bruxelles, CESifo and CEPR
Alexander Sebald, University of Copenhagen

Discussion Paper No. 7365
July 2009

Centre for Economic Policy Research
53–56 Gt Sutton St, London EC1V 0DG, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 7365

July 2009

ABSTRACT

Decision-making Procedures: A General Theory and Its Field Experimental Test

It is a persistent finding in psychology and experimental economics that people's behavior is not only shaped by outcomes but also by decision-making procedures. In this paper we develop a general framework capable of modelling these procedural concerns. Within the context of psychological games we define procedures as mechanisms that influence the probabilities of reaching different endnodes. We show that for such procedural games a sequential psychological equilibrium always exists. Applying this approach within a principal-agent context we show that the way less attractive jobs are allocated is crucial for the effort exerted by agents. This prediction is tested in a field experiment, where some subjects had to type in data, whereas others had to verify the data inserted by the typists. The controllers' wage was 50% higher than that of the typists. In one treatment the less attractive typists' jobs were allocated directly, whereas in the other treatment the allocation was done randomly. As predicted, random allocation led to higher effort levels of the typists than direct appointment.

JEL Classification: A13, C70, C93 and D63
Keywords: appointment procedures, procedural concerns and psychological game theory

Gani Aldashev
FUNDP
Rue de Bruxelles 61
5000 Namur
BELGIUM

Email: gani.aldashev@fundp.ac.be

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=162714

Georg Kirchsteiger
ECARES
Université Libre de Bruxelles
Avenue F D Roosevelt 50
1050 Brussels
BELGIUM

Email: gkirchst@ulb.ac.be

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=132740

Alexander Sebald
Dept of Economics
University of Copenhagen
Øster Farimagsgade 5
1353 Copenhagen
DENMARK

Email: alexander.sebald@econ.ku.dk

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=164632

Submitted 07 July 2009

We are very grateful to Martin Dufwenberg, Estelle Cantillon, Paolo Casini, Paola Conconi, Werner Güth and seminar participants at the University of Konstanz.

1 Introduction

Among psychologists a broad consensus exists that not only expected outcomes shape human behavior, but also the way in which decisions are taken. A persistent finding is that people behave very differently in outcomewise-identical situations depending on the decision making procedures which led to these outcomes [e.g. Thibaut and Walker (1975), Lind and Tyler (1988), Collie et al. (2002), Anderson and Otto (2003) and Blader and Tyler (2003)]. Workplace relations is a prominent example of an economic context in which procedures have been found to play an eminent role. Reactions to promotion decisions, bonus allocations, dismissals etc. strongly depend on the perceived fairness of selection procedures [e.g. Lemons and Jones (2001), Konovsky (2000), Bies and Tyler (1993), Lind et al. (2000) and Roberts and Markel (2001)]. In this literature it is argued that procedures matter because they affect the beliefs that people hold about each others' intentions and expectations which subsequently influence their behavior.

But not only psychologists have found evidence for procedural concerns. There is also evidence from experimental economics which indicates that people are not only concerned about the consequences of decisions, but also care about the decision-making procedures involved [Blount (1995), Falk et al (2000), Bolton et al. (2005), Charness (2004), Brandts et al. (2006), Charness and Levine (2007)]. Brandts et al. (2006), for example, show that selection procedures matter in a three-player game in which one player has to select one of the other players to perform a specific task.[1] They find that the selected players behave very differently in their subsequent tasks depending on the procedure which was used to select them.

Traditionally, economic theory assumes that agents only care about the consequences of decisions. Although consequentialism allows that agents care about the payoffs of other players, e.g. that they are altruistic or envious, it inherently implies that people behave the same in outcomewise identical situations, regardless of the decision-making procedures that led to these situations. Thus, consequentialism is at odds with the aforementioned evidence that people are (also) concerned about the way in which decisions are taken. To get an intuition for why consequentialism fails when procedures matter, consider the following principal-agent relation: Suppose there is a profit-maximizing principal that has to assign two equally skilled agents to two different jobs. The first job ('controller') is paid better than the second ('typist'), with no difference in difficulty. For simplicity, let's only concentrate on the disadvantaged, i.e. the typist, and his effort choice. Assume the principal first chooses the procedure by which the typist is chosen. He has three different strategies: (i) He can directly appoint agent 1; (ii) he can directly appoint agent 2; (iii) he can choose a random appointment procedure giving both agents an equal chance to get either job. After the tasks are allocated, the appointed typist

[1] Two different treatments are studied which differ with regard to the selection procedure. In both treatments the task of the selected player is to choose between the same two payoff allocations determining the payoffs of all three players.

chooses his effort. Note that the principal basically has two appointment procedures that he can choose between. He can either decide himself by using strategies (i) and (ii) or he can let chance decide by using strategy (iii). These appointment procedures differ with regard to the ex-ante probabilities that they attach to specific outcomes. The procedure implied by strategies (i) and (ii) puts probability 1 on one of the agents. Procedure (iii), on the other hand, puts the ex-ante probability 0.5 on each of the agents. Clearly, the typist's effort choice should be independent of the principal's selection procedure, if the typist is only concerned about final outcomes. His effort choice should be the same irrespective of the selection procedure, as the feasible final payoff allocations are the same independent of how the typist was chosen.

In contrast to this, however, the aforementioned behavioral evidence from psychology and experimental economics suggests that the typist's effort choice depends on the selection procedure even though the feasible final payoff allocations are independent of how he was chosen. More precisely, the typist's effort choice should be higher when the principal uses the unbiased random assignment procedure (i.e. strategy (iii)). In line with this behavioral evidence Sebald (2009) suggests that 'procedural concerns' can be conceptualized by assuming that people have belief-dependent reciprocal preferences à la Dufwenberg and Kirchsteiger (2004). In the principal-agent context the argument runs as follows: Imagine that agents are not only concerned about final outcomes, but also care about the kindness and unkindness of the principal towards them. The typist's perception about the principal's (un)kindness towards him depends on the procedure that the principal uses to take his appointment decision. If the principal decides to choose the typist directly (i.e. strategies (i) or (ii)), the chosen agent interprets the principal's decision as intentionally directed against him. If, on the other hand, the principal decides to use a random appointment procedure (i.e. strategy (iii)), the agent interprets the outcome as pure chance rather than an intentional act of the principal. The chosen agent interprets the principal's choice of the random appointment procedure as 'kinder' and subsequently exerts higher effort compared to the situation in which he is directly chosen to be the typist.

Our principal-agent setting and Sebald (2009)'s argument highlight two important aspects of procedures and procedural concerns. First, procedures can be viewed as possibly stochastic decision-making mechanisms determining the ex-ante probabilities for situations in which agents can find themselves in ex-post. Second, procedural concerns mean that these ex-ante probabilities have an impact on agents' decisions even ex-post, i.e. after the resolution of the uncertainty inherent in decision-making procedures. Following this intuition, agents that exhibit procedural concerns are not consequentialist, for them 'bygones are not bygones'.

It has been shown that, in order to model such a non-consequentialist behavior, one has to introduce belief-dependent utilities, i.e. one has to use the concepts of psychological game theory [see Geanakoplos et al. (1989) and Battigalli and Dufwenberg (2009)]. Following Sebald (2009), we intuitively explained procedural concerns using reciprocity concerns in the example above [see Rabin (1992) and Dufwenberg and Kirchsteiger (2004)]. But reciprocity is just one possible belief-dependent concern. Other types of belief-dependent emotions (e.g. regret, disappointment, guilt) have already been formalized using psychological game theory. For example, Charness and Dufwenberg (2006) as well as Battigalli and Dufwenberg (2007) look at the interaction of agents that are guilt averse. Ruffle (1999) presents a psychological game-theoretic model in which surprise, disappointment and embarrassment arise in the interaction of emotional agents. Battigalli and Dufwenberg (2009) present a general framework for dynamic psychological games encompassing all different types of belief-dependent preferences.

These papers develop models featuring different belief-dependent preferences and investigate their impact on strategic interactions. However, none of them deals with the role of procedural choices. In contrast, this paper concentrates on procedures, i.e. stochastic decision-making mechanisms, and the impact of procedural choices on the strategic interaction of agents that are (also) motivated by belief-dependent preferences. More precisely, this paper integrates stochastic decision-making procedures into a general model of belief-dependent preferences, and thus provides a framework which allows for the analysis of strategic situations in which people

(i) are motivated by belief-dependent preferences and

(ii) can choose between possibly stochastic decision-making mechanism.

Our general framework is based on the theory of dynamic psychological games by Battigalli and Dufwenberg (2009). Unlike their model, we allow for moves of chance in order to allow for procedural concerns.[2] Crucially, our general framework shows that procedural concerns are not only an artifact of reciprocal preferences as suggested by Sebald (2009), but that belief-dependent preferences in general (e.g. guilt, regret and disappointment) might imply procedural concerns. Using our framework we formalize the principal-agent example sketched above and derive a hypothesis concerning the impact of the principal's procedural choice on the typist's effort decision.

In addition to our general theoretical framework, the paper's second key contribution is the test of our theory in a field experiment. To the best of our knowledge, the evidence for procedural concerns have not been yet found in a field-experimental setting. Our experimental findings thus have the advantage that they are established in the setting where the subjects operated in their natural environment, that they did not know that they participated in an experiment, and that their effort choices were real (and not abstract numbers). We hired undergraduate students as research assistants for an ongoing data-building project. The task of one set of subjects was typing in data, while that of another set was verifying the data inserted by the typists. The controllers' wage was

[2]Without reference to procedural concerns Battigalli and Dufwenberg (2009) already suggest (pp 26-28) that their framework might also be extended to allow for moves of chance.

50% higher than that of the typists. The experiment had two treatments. In the first treatment, we allocated subjects to jobs directly, whereas in the second treatment, the allocation was explicitly randomized.

Our findings support the theoretical hypothesis derived on the basis of our principal-agent setting. Furthermore, we establish a novel result concerning gender differences in procedural concerns. Under random allocation, male typists exerted significantly higher effort than under direct appointment, both in terms of quantity and quality. Female typists, on the other hand, exerted higher effort only in terms of quality. These findings relate to and complement the existing literature on gender differences in social preferences (see Croson and Gneezy 2009).

The organization of the paper is as follows: In the next section the basic setup with belief-dependent utilities and procedures is described. In section 3 we extend the concept of sequential psychological equilibria to our framework, and prove its existence. We present a formal version of the principal-agent setting sketched above and derive hypotheses regarding the impact of the principal's procedural choices in section 4. In section 5 we describe the field experiment and report its results. In section 6 we conclude.

2 Belief-Dependent Utilities and Procedures

In this section we define the class of extensive form games with belief-dependent payoffs and moves of chance that we use to model procedural concerns. As said before, belief-dependent incentives are necessary to formalize concerns like reciprocity and guilt aversion, while moves of chance are needed to model procedures.

Let the set of players be $\mathcal{N} = \{0, 1, ..., N\}$ where 0 denotes the uninterested player *chance*. Denote as $\mathcal{H}$ the finite set of histories h, with the empty sequence $h^0 \in \mathcal{H}$, and Z the set of end-nodes. Histories $h \in \mathcal{H}$ are sequences that describe the choices that players have made on the path to history h. At each non-terminal history each player $i \in \mathcal{N}$ disposes a nonempty, finite sets of feasible actions $\mathcal{A}_i(h)$. After each non-terminal history $h \in \mathcal{H} \backslash Z$ every player $i \in \mathcal{N} \backslash \{0\}$ chooses an action $a_i(h)$ from $\mathcal{A}_i(h)$. Player 0 "chooses" an action $a_0(h)$ from $\mathcal{A}_0(h)$ according to a commonly known and verifiable probability distribution $\omega(h)$ defined on $\mathcal{A}_0(h)$ with full support. Note $\mathcal{A}_i(h)$ can be a singleton, meaning that player i is inactive at history h. Therefore, our framework with moves of chance also allows for games with perfect information.

Let $\mathcal{A}_i$ be the finite set of pure strategies of players $i \in \mathcal{N} \backslash \{0\}$ and ω the single verifiable behavioral strategy of player 0. Pure strategies of players $i \in \mathcal{N}$ are denoted by $a_i = (a_i(h))_{h \in \mathcal{H} \backslash Z}$. Note that we use pure strategies of player 0 just for notational purposes. Each of them just constitutes the sequence of realizations of the commonly known behavioral strategy $\omega = (\omega(h))_{h \in \mathcal{H} \backslash Z}$.

To formalize belief-dependent payoffs we assume that players hold

(i) beliefs about the strategies of other players,

(ii) beliefs about the beliefs of other players,

(iii) beliefs about the beliefs about the beliefs of other players etc,

i.e. we assume that players have an infinite hierarchy of beliefs, and we

(iv) let them update their beliefs as events unfold.

In order to formally capture assumptions (i)-(iv), we have to define an epistemic structure (*collectively coherent hierarchies of beliefs*) which describes what people initially belief and how they update their beliefs as play unfolds. This epistemic structure can be characterized in the context of our class of extensive form games by assuming that players hold hierarchies of conditional beliefs over the strategies as well as beliefs of other players $i \in \mathcal{N} \setminus \{0\}$.

We only summarize the theory of hierarchies of conditional beliefs.[3] We describe, first, a system of conditional first-order beliefs and, second, show how this extends to higher-orders (i.e. second-order beliefs etc). Denote by $\mathcal{A}_{-i}$ the set of pure strategies combinations of all the opponents j with $j \in \mathcal{N} \setminus \{0, i\}$. At the beginning of any game, i.e. at the initial history h^0, player $i \neq 0$ does not know the true strategies of his opponents. He only learns the true strategy $a_{-i} \in \mathcal{A}_{-i}$ step-by-step by updating his beliefs as the game unfolds. More formally, player i assigns probabilities to the events in the Borrel sigma algebra $\mathcal{B}$ of $\mathcal{A}_{-i}$ according to some probability measure. Let $\Delta(\mathcal{A}_{-i})$ be the set of all such probability measures. Denote $\mathcal{C} \subseteq \mathcal{B}$ the set of potential conditioning events at which player i can update his beliefs. In other words, $\mathcal{C}$ is the set of potentially observable events. Player i holds probabilistic beliefs about his opponents's pure strategies conditional on each event $F \in \mathcal{C}$. These probabilistic beliefs are captured in a conditional probability system (*cps*).

Following Battigalli and Dufwenberg (2009) consider the following definition:

Definition 1 *A conditional probability system (cps) is a function* $\mu(\cdot|\cdot) : \mathcal{B} \times \mathcal{X} \rightarrow [0, 1]$ *defined on* $(X, \mathcal{B}, \mathcal{C})$ *such that for all* $E \in \mathcal{B}$ *and* $F', F \in \mathcal{C}$*:*

1. $\mu(\cdot|\cdot) \in \Delta(X)$,

2. $\mu(F|F) = 1$,

3. $E \subseteq F' \subseteq F$ *implies* $\mu(E|F) = \mu(E|F')\,\mu(F'|F)$,

where X *is a set, e.g.* $\mathcal{A}_{-i}$*, whose 'true' element* $x \in X$ *is initially unknown and only learned step-by-step as conditioning events, e.g.* $F \in \mathcal{C}$*, are reached.*

[3]For topological details, proofs and further references see Brandenburger and Dekel (1993) and Battigalli and Siniscalchi (1999).

First order beliefs mean that $X = \mathcal{A}_{-i}$. The first two conditions of definition 1 ensure that $\mu(\cdot|F)$ is indeed a probability measure (i.e. $\mu(\cdot|F) \in \Delta(X)$) which puts all probability weight on F given that F is observed. Condition 3 ensures that players update their beliefs according to Bayes' rule. The set of all functions μ for which conditions *1-3* hold is denoted by $\Delta^H(X)$. Hence, $\Delta^H(\mathcal{A}_{-i})$ is the set of all conditional probability systems of order 1 of player i.

Definition 1 can easily be extended to higher-order beliefs. In the construction of the first-order *cps* we start from an initial situation in which player i does not know the true pure strategy of his opponents. He has a conditional first-order belief over it which is updated as play unfolds. Analog to this, in the construction of a second-order belief we start from an initial situation in which player i does not know the true pure strategy and the true conditional first-order belief of players $-i$. Hence, the relevant set X in definition 1 becomes:

$$X = \mathcal{A}_{-i} \times \prod_{j \neq i} \Delta^H(\mathcal{A}_{-j}),$$

where $i, j \in \mathcal{N} \backslash \{0\}$ and $\Delta^H(\mathcal{A}_{-j})$ is the set of conditional first-order *cps* of player j. The resulting conditional probability system does not only represent player i's belief about the strategy of players $-i$, but also about their first-order beliefs.

Generalizing this idea, first- and higher-order *cps* are defined recursively as follows. Let:

$$\begin{aligned} X_{-i}^{0} &= \mathcal{A}_{-i}, \text{ where } i \in \mathcal{N} \backslash \{0\}, \\ X_{-i}^{k} &= X_{-i}^{k-1} \times \prod_{j \neq i} \Delta^H\left(X_{-j}^{k-1}\right), \text{ where } i \in \mathcal{N} \backslash \{0\} \text{ and } k = 1, 2, \ldots . \end{aligned}$$

Then, a *cps* $\mu_i^k \in \Delta^H\left(X_{-i}^{k-1}\right)$ is called a k-order *cps* or simply a k-order belief. For $k > 1$, μ_i^k is a joint *cps* on the opponents' strategies and $(k-1)$-order *cps*', i.e.:

$$\begin{aligned} \mu_i^1 &\in \Delta^H\left(X_{-i}^0\right) \text{ where } X_{-i}^0 = \mathcal{A}_{-i}, \\ \mu_i^2 &\in \Delta^H\left(X_{-i}^1\right) \text{ where } X_{-i}^1 = \mathcal{A}_{-i} \times \Delta^H(\mathcal{A}_{-j}), \\ \mu_i^3 &\in \Delta^H\left(X_{-i}^2\right) \text{ where } X_{-i}^2 = \mathcal{A}_{-i} \times \Delta^H(\mathcal{A}_{-j}) \times \Delta^H\left(\mathcal{A}_{-j} \times \Delta^H(\mathcal{A}_{-i})\right) \text{ etc. } . \end{aligned}$$

This brings us to the formal definition of hierarchies of *cps*':[4]

Definition 2 *A hierarchy of cps is a countably infinite sequence of cps':*

$$\mu_i = \left(\mu_i^1, \mu_i^2, \ldots\right) \in \prod_{k>0} \Delta^H\left(X_{-i}^{k-1}\right).$$

[4]See also Battigalli and Dufwenberg (2009), p. 11.

As one can see, each piece of information appears many times in the belief hierarchy of player i. This implies that one can calculate marginal beliefs of higher-order beliefs. As also Geanakoplos et al. (1989) point out, these marginal beliefs of higher-order beliefs should coincide with lower-order beliefs in the belief hierarchy for the hierarchy to be meaningful. In other words beliefs should be *coherent*. We say a hierarchy of *cps*' is coherent if the *cps*' of distinct orders assign the same conditional probabilities to lower-order events. This means,

$$\mu_i^k(\cdot|h) = marg_{X_{-i}^{k-1}} \mu_i^{k+1}(\cdot|h)\,(k = 1, 2, ...; h \in \mathcal{H}),$$

where $marg_{X_{-i}^{k-1}} \mu_i^{k+1}(\cdot|h)$ is the event of order $k-1$ in the *cps* of order $k+1$, $\mu_i^{k+1}(\cdot|h)$. If this condition holds, player i is said to have a coherent conditional belief system. It can be shown that a coherent hierarchy of *cps'* induces a single *cps* ν_i on the cross product of $\mathcal{A}_{-i}$ and the sets of hierarchies of *cps*' of i's opponents $-i$. Note, however, coherency regarding the own beliefs does not exclude the possibility that the *cps* ν_i puts a positive probability on the opponents incoherence. But as players are rational they should not believe that their opponents entertain incoherent beliefs. Hence, in order to rule this out, say that a coherent hierarchy μ_i satisfies belief in coherency of order 1 if the induced *cps* ν_i is such that each $\nu_i(\cdot|h)$ with $h \in \mathcal{H}$ assigns probability one to the opponents' coherence of order 1. The hierarchy of coherent beliefs μ_i satisfies belief in coherency of order k, if it satisfies belief in coherency of order $k-1$, μ_i is *collectively coherent*, if it satisfies belief in coherency of order k for each positive integer k.[5] We denote the set of *collectively coherent hierarchies of beliefs* of player i by M_i. The set of collectively coherent beliefs of the opponents $-i$ is M_{-i} and $M = \prod_{j \in \mathcal{N}\backslash\{0\}} M_j$.

As the probability distributions associated with the moves of the player *chance*, i.e. player 0, are commonly known, nobody faces any uncertainty with regard to his true strategy ω. Players do not learn the true strategy of player 0 over the course of the game since it is common knowledge from the beginning anyhow.

Finally, belief-dependent utilities are defined as:

Definition 3 *A belief dependent utility u of player $i \in \mathcal{N}\backslash\{0\}$ is a function:*

$$u_i : \mathcal{Z} \times \{\omega\} \times \mathcal{M}_i \times \Pi_{j \neq i}(\mathcal{A}_j \times \mathcal{M}_j) \to \Re,$$

where $i, j \in \mathcal{N}\backslash\{0\}$.

Hence, as in Battigalli and Dufwenberg (2009), the utility of any player i depends on the end-node, $\mathcal{Z}$, the hierarchy of conditional beliefs of player i and the hierarchies of conditional beliefs of all other players $j \neq i$. In addition to this, however, we also assume that player i's utility depends on the verifiable behavioral strategy of the player chance,

[5] See also Battigalli and Dufwenberg (2009)

ω. The reason for this is the following: Different to the strategies of all players $j \neq i$, the strategy of the player chance is not slowly unraveled in the course of the game, i.e. it is not 'learned'. It is commonly known and it is known that it is commonly known. Hence, players do not update their beliefs concerning the moves of chance. Given this, we assume that players' utilities do not depend on the hierarchy of conditional beliefs of player 0, but simply on his verifiable strategy ω.

As also argued in the introduction, allowing for moves of chance and assuming that players' utilities depend on the verifiable behavioral strategy of the player 0 allows (i) to capture the behavioral evidence on procedural concerns and (ii) to analyze the impact of procedural choices on the strategic interaction of agents that are motivated by belief-dependent preferences. As also exemplified in our introductory principal-agent story, procedures can be viewed as possibly stochastic decision mechanisms determining the ex-ante probabilities for situations in which agents can find themselves in ex-post. Remember, in our principal-agent setting there are basically two types of decision-making procedures that the principal can use to take his decision. On the one hand, the principal can decide himself using strategies (i) and (ii) which both put probability 1 on one of the agents. On the other hand, he can use a decision making procedure giving each agent a verifiable chance of 50% to get either job by using strategy (iii). This exemplifies, how decision-making procedures can be conceptualize by allowing for moves of chance. Or in other words, by allowing for moves of chance we can formalize strategic envirmonments in which people have the possibility to choose between different decision making procedures.

How do these procedural choices impact strategic interactions? As already argued in the introduction, if people are consequentialists, their behavior is independent of the decision-making procedures involved. This is at odds with the behavioral evidence from psychological and economic research. The persistent evidence on procedural concerns can be rationalized, however, assuming that agents have belief-dependent preferences as formalized in the general framework in this section. To get an intuition for how our framework explains the empirical evidence on procedural concerns, consider the following example: Assume the principal in our principal-agent setting is guilt averse á la Battigalli and Dufwenberg (2007). Note, guilt aversion á la Battigalli and Dufwenberg (2007) is one type of belief-dependent emotion that is in line with definition 3. In the context of our principal-agent example this means, the principal is averse to not living up to his belief about the agents' expectations about his decision. More formally, the principal is assumed to have the following utility function:

$$u_p = x_p + \Pi_{j \neq p} G_{pa} \cdot (x_a - E_{pa}[x_a]) \tag{1}$$

where x_p and x_a respectively are the material payoffs of the principal and the agent, G_{pa} is an exogenous constant defining the principal's sensitivity to guilt towards agent a and $E_{pa}[x_a]$ is the principal's belief about the agent's payoff expectation.[6] Hence,

[6]For a more detailed description of guilt aversion see Battigalli and Dufwenberg (2007)

the principal's reduction of utility from his feeling of guilt towards agent a is given by $G_{pa}\cdot(x_a - E_{pa}[x_a])$. As one can easily see, the principal feels guilt, i.e. $G_{pa}\cdot(x_a - E_{pa}[x_a]) < 0$, if he does not live up to the expectations of the agent, i.e. $x_a < E_{pa}[x_a]$.

Assume now, the principal believes the agents believe that he will use the impartial random appointment procedure which gives both of them an equal chance to get the higher and the lower paid job. More precisely, the principal believes that the agents' expectation about their payoff is

$$0.5 \cdot w_c + 0.5 \cdot w_t,$$

where 0.5 is the chance associated to the random appointment procedure and w_c and w_t are respectively the wages associated with the higher and the lower paying job. Given this, the principal's guilt associated with each type of decision-making procedure that he can use to take his decision can be described as follows. First, if the principal, given his beliefs, chooses to decide himself, his guilt towards the typist would be equal to

$$G_{pt} \cdot (w_t - [0.5 \cdot w_c + 0.5 \cdot w_t]) = G_{pt} \cdot (-0.5 \cdot w_c + 0.5 \cdot w_t).$$

As $w_c > w_t$, $G_{pt} \cdot (-0.5 \cdot w_c + 0.5 \cdot w_t) < 0$, i.e. guilt would reduce his utility. Second, if he decides to take the random decision-making procedure and the same agent becomes typist, his feeling of guilt towards him is:

$$G_{pt} \cdot ([0.5 \cdot w_c + 0.5 \cdot w_t] - [0.5 \cdot w_c + 0.5 \cdot w_t]) = 0.$$

The principal's feeling of guilt would be 0 following the random decision-making procedure as he would life up to his belief about the agents' expectations.

Depending on the decision-making procedure that the principal uses to take his appointment decision, he feels more or less guilt. Quite intuitively, the different levels of guilt associated with the different decision-making procedures might influence the principal's procedural choice. In other words, the principal is concerned about the procedure used as it influences his feeling of guilt towards the agents. Obviously guilt aversion, is just one example of a belief-dependent motivation. However, this example shows how moves of chance can be used to formalize procedures and how the stochastic character underlying these decision-making procedures influences the evaluation of guilt.

Our general formulation in equation 3 allows for all kinds of belief-dependent preferences (e.g. reciprocity, disappointment, regret). More technically, our class of extensive form games basically represents an extension of Battigalli and Dufwenberg (2009)'s class of extensive form games with belief-dependent preferences allowing for moves of chance. This means our framework with moves of chance allows for the analysis of procedural concerns in strategic interactions in which agents have all kinds of belief-dependent psychological motivations.

In the next section we present a solution concept which can be used to derive predictions regarding the strategic interaction of agents that are motivated by belief-dependent preferences in situations in which outcomes also depend on moves of chance.

3 Sequential Equilibria

Battigalli and Dufwenberg (2009) adapt Kreps and Wilson (1982)'s concept of sequential equilibrium to their class of dynamic psychological games. They do so by characterizing *consistent assessments* that do not only consist of first-, but also of higher-order beliefs and defining sequential equilibria as sequentially rational consistent assessments.

As in Battigalli and Dufwenberg (2009), also our equilibrium concept refers to mixed strategies, i.e. implicit randomizations over sets of procedures. Note, however, that, following Aumann and Brandenburger (1995), we interpret player i's mixed strategy as a conjecture on the part of his opponents as to what player i will do. Hence, denote a behavioral strategy of player i as $\sigma_i = (\sigma_{i,h})_{h\in\mathcal{H}_i} \in \Sigma_i$, where Σ_i is the set of all mixed strategies of player i. The behavioral choice $\sigma_{i,h} \in \Sigma_i(h)$ in h has to be understood as an implicit randomization over the set of actions $\mathcal{A}_i(h)$ in history h and interpreted as an array of common conditional first-order beliefs held by i's opponents.[7] This means that the behavioral strategy σ_i is part of an assessment $((\sigma,\rho_0),(\mu,\rho_0)) = ((\sigma_i,\omega),(\mu_i,\omega))_{i\in\mathcal{N}\setminus\{0\}}$ of behavioral strategies and hierarchies of conditional beliefs.

Three conditions ensure consistency of assessments in the original characterization by Kreps and Wilson (1982):

1. Beliefs must be derived using Bayes' rule,
2. Beliefs must reflect that players choose their strategies independently,
3. Players with identical information have identical beliefs.

In addition to these conditions, Battigalli and Dufwenberg (2007) add another requirement for consistency:

4. Players hold correct beliefs about each others' beliefs.

Condition 1 holds by the definition of hierarchies of conditional belief systems (Definition 1). In other words, hierarchies of beliefs are defined in such a way that conditional beliefs are consistent with Bayes' rule. In order to formalize conditions 2-4 we first need to define what is meant by *stochastic independence*. Note, the observability of past actions allows us to define stochastic independence of the conditional belief systems in terms of marginal *cps*'. Different to the concept of marginal beliefs used in the previous section, a marginal *cps* here refers to player i's marginal belief on the strategies of a particular player j and it is denoted as $\mu^1_{ij} \in \Delta^H(\mathcal{A}_j)$, where $\Delta^H(\mathcal{A}_j)$ is the set of marginal *cps* on the strategies of player j. Given this we can define stochastic independence of beliefs as:[8]

[7] See Battigalli and Dufwenberg (2009)

[8] See also Battigalli and Dufwenberg (2009)'s definition of stochastic independence

Definition 4 *A first-order cps $\mu_i^1 \in \Delta^H(\mathcal{A}_{-i})$ satisfies stochastic independence, if there exists a profile of marginal cps' $(\mu_{ij}^1)_{j\neq i} \in \prod_{j\neq i} \Delta^H(\mathcal{A}_j)$ such that*

$$\mu_i^1(\mathcal{A}_{-i}|h) = \prod_{j\neq i} \mu_{ij}^1(\mathcal{A}_j|h)$$

for all $h \in \mathcal{H}_i$. We denote the set of stochastically independent first-order cps' of a player i as $\Delta_I^H(\mathcal{A}_{-i})$.

This brings us to our definition of consistent assessments:

Definition 5 *An assessment $((\sigma,\omega),(\mu,\omega))$ is consistent if:*

1. *The first-order cps of each player satisfies stochastic independence as formalized in Definition (4), i.e.:*
$$\forall i \in \mathcal{N}\setminus\{0\}, \mu_i^1 \in \Delta_I^H(\mathcal{A}_{-i}).$$
2. *The marginal first-order cps of any two players about any third player coincide, i.e.:*
$$\forall i \in \mathcal{N}\setminus\{0\}, \forall l \in \mathcal{N}\setminus\{i,j,0\}, \forall h \in \mathcal{H}, \mu_{il}^1(\cdot|h) = \mu_{jl}^1(\cdot|h).$$
3. *Each players higher order beliefs in μ assign probability 1 to the lower order beliefs in μ itself:*
$$\forall i \in \mathcal{N}\setminus\{0\}, \forall k > 1, \forall h \in \mathcal{H}, \mu_i^k(\cdot|h) = \mu_i^{k-1}(\cdot|h) \times \delta_{\mu_{-i}^{k-1}},$$
where $\delta_{\mu_{-i}^{k-1}}$ is the probability measure which assigns probability 1 to μ_{-i}^{k-1}.

Conditions 1 and 2 capture the assumption that beliefs should be the end-product of a transparent reasoning process of intelligent people [Battigalli and Dufwenberg (2009)]. Condition 3, on the other hand, is analog to Geanakoplos et al. (1989)'s condition requiring that players hold common and correct beliefs about each others' beliefs.

After having defined consistent assessments we can formally characterize *sequential psychological equilibria* (henceforth: SPE) by requiring sequential rationality:

Definition 6 *An assessment $((\sigma,\omega),(\mu,\omega))$ is a sequential psychological equilibrium (SPE), if for all $i \in \mathcal{N}\setminus\{0\}, h \in \mathcal{H}_i$ it holds:*

$$Supp(\sigma_{i,h}) \subseteq argmax_{\mathcal{A}_{i,h}\in\mathcal{A}(h)} E_{\mu,\omega}[u_i|h,\mathcal{A}_{i,h}],$$

where $E_{\mu,\omega}[u_i|h,\mathcal{A}_{i,h}]$ is the expected utility of player i conditional on history h, choice $\mathcal{A}_{i,h} \in \mathcal{A}(h)$ and given the system of hierarchies of conditional beliefs μ and the commonly known strategy, ω, played by player 0.

Note, the expected utility of any player $i \in \mathcal{N} \setminus \{0\}$ (conditional on history h, choice $\mathcal{A}_{i,h} \in \mathcal{A}(h)$, given the system of consistent hierarchies of conditional beliefs μ and the commonly known strategy, ω) can be defined as:

$$E_{\mu,\omega}\left[u_i | h, \mathcal{A}_{i,h}\right] = \sum_{a_0 \in \mathcal{A}_0(h)} \omega\left(a_0 | h\right) \sum_{\mathcal{A}_{-i} \in \mathcal{A}_{-i}(h)} \mu_i^1\left(\mathcal{A}_{-i} | h\right)$$
$$\sum_{\mathcal{A}_i \in \mathcal{A}_i\left(h, \mathcal{A}_{i,h}\right)} \mu_{ji}^1\left(\mathcal{A}_i | \left(h, \mathcal{A}_{i,h}, \mathcal{A}_{-i,h}\right)\right) u_i\left(\zeta\left(\mathcal{A}_i, \mathcal{A}_{-i}, a_0\right), \omega, \mu, \mathcal{A}_{-i}\right),$$

where $\zeta\left(\mathcal{A}_i, \mathcal{A}_{-i}, a_0\right) \in Z$ denotes the terminal history induced by the strategies $\mathcal{A}_i$ and $\mathcal{A}_{-i}$, and the strategy a_0 of player 0. Note, this specification is different from the expected utility formula traditionally used. Furthermore, it is also different from the specification used by Battigalli and Dufwenberg (2009) as it encloses the behavioral moves of the player chance.

The following proposition shows that there exists at least one sequential psychological equilibrium in any game with psychological incentives and continuous utility functions:

Proposition 1 *If the utility functions are continuous, there exists at least one sequential psychological equilibrium assessment.*

Proof: Consider a game with psychological incentives in which any action at any history is played with a strictly positive minimal probability ε. More formally, consider an ε-perturbed game Γ^ε in which players $i \in \mathcal{N} \setminus \{0\}$ dispose of 'constrained' choice sets $\Sigma_i^\varepsilon(h)$ at each history $h \in \mathcal{H}_i$. The 'constrained' choice set $\Sigma_i^\varepsilon(h)$ of player i in history h is defined as:

$$\Sigma_i^\varepsilon(h) := \{\tau_{i,h} \in \Sigma_i(h) | \tau_{i,h}(\mathcal{A}_{i,h}) \geq \varepsilon, \forall \mathcal{A}_{i,h} \in \mathcal{A}_i(h)\}.$$

So $\Sigma_i^\varepsilon(h)$ consists of only those elements in $\Sigma_i(h)$ that put a strictly positive probability greater or equal to ε on all elements $\mathcal{A}_{i,h} \in \mathcal{A}_i(h)$, i.e. $\Sigma_i^\varepsilon(h) \subset \Sigma_i(h)$. It follows that in any Γ^ε the set of strictly mixed procedural strategies of players $i \in \mathcal{N} \setminus \{0\}$ is $\Sigma_i^\varepsilon = \times_{h \in \mathcal{H}_i} {\Sigma_i}^\varepsilon(h)$ and the set of all strictly positive behavioral strategy profiles is $\Sigma^\varepsilon := \times_{i \in \mathcal{N} \setminus \{0\}} \Sigma_i^\varepsilon$. Note, for each $\sigma \in \Sigma^\varepsilon$ there exists a unique corresponding profile of hierarchies of *cps*' $\mu = \beta(\sigma)$ such that $((\sigma, \omega), (\beta(\sigma), \omega))$ is consistent.

Now, define for $\sigma \in \Sigma^\varepsilon$, $\varepsilon > 0$, $i \in \mathcal{N} \setminus \{0\}$ and $h \in \mathcal{H}_i$ the local best-response of player i in history h as:

$$BR_{i,h}^\varepsilon(\sigma) := \{\hat{\tau}_{i,h} \in \Sigma_i^\varepsilon(h) | u_i(\sigma_i / \hat{\tau}_{i,h}, \sigma_{-i}, \omega) \geq u_i(\sigma_i / \tau_{i,h}, \sigma_{-i}, \omega), \forall \tau_{i,h} \in \Sigma_i^\varepsilon(h)\},$$

where $\sigma_i / \tau_{i,h}$ denotes the behavioral strategy for player i that specifies the strictly positive mixture $\tau_{i,h}$ at history $h \in \mathcal{H}_i$ and σ_i at every other history controlled by player i. In other words, local best-response-correspondences are strictly mixed behavioral choices that put at least a minimum probability ε on each procedure $\mathcal{A}_{i,h} \in \mathcal{A}_i(h)$ given i's choices in all other histories controlled by him and given the behavioral strategy of the opponents. The domain of the local best-response-correspondence is Σ^{ε}. The set $\Sigma^{\varepsilon}=\Sigma_1^{\varepsilon}\times\Sigma_2^{\varepsilon}...\times\Sigma_N^{\varepsilon}$ and each Σ_i^{ε} with $i\in\mathcal{N}\backslash\{0\}$ is defined as $\Sigma_i^{\varepsilon}=\times_{h\in\mathcal{H}}\Sigma_i^{\varepsilon}(h)$. As said above, $\Sigma_i^{\varepsilon}(h)$ is the set of all behavioral strategies of player i at history h that put at least a strictly positive probability ε on each action $\mathcal{A}_{i,h}\in\mathcal{A}_i(h)$. It is non-empty (because $\mathcal{A}_i(h)$ is non-empty), compact and convex. Hence, also Σ^{ε} is non-empty, compact and convex (because the finite Cartesian product of nonempty, convex and compact sets is itself nonempty, convex and compact). Furthermore, $BR_{i,h}^{\varepsilon}(\sigma)$ is upper-semi-continuous. Note, the local best-response-correspondence $BR_{i,h}^{\varepsilon}(\sigma)$ is upper-semi-continuous, if for any sequence $(\sigma_i/\hat{\tau}_{i,h}^m,\sigma_{-i}^m)\rightarrow(\sigma_i/\hat{\tau}_{i,h},\sigma_{-i})$ such that $\sigma_i/\hat{\tau}_{i,h}^m\in BR_{i,h}^{\varepsilon}(\sigma_i/\hat{\tau}_{i,h}^m,\sigma_{-i}^m)$ for all $m\in\{1,2,...\}$, we have $\sigma_i/\hat{\tau}_{i,h}\in BR_{i,h}^{\varepsilon}(\sigma_i/\hat{\tau}_{i,h},\sigma_{-i})$. To see that this is indeed the case, note that for all m, the $u(\sigma_i/\hat{\tau}_{i,h}^m,\sigma_{-i}^m)\geq u(\sigma_i/\hat{\tau}_{i,h}',\sigma_{-i}^m)$ for all $\sigma_i/\hat{\tau}_{i,h}'\in\Sigma_i^{\varepsilon}$. Hence, by the continuity of the utility function, we have $u(\sigma_i/\hat{\tau}_{i,h},\sigma_{-i})\geq u(\sigma_i/\hat{\tau}_{i,h}',\sigma_{-i})$.

Given the local best-response correspondence $BR_{i,h}^{\varepsilon}(\sigma)$, the best-response correspondence $BR^{\varepsilon}(\sigma)$ is defined as:

$$BR^{\varepsilon}=\left(\hat{\tau}_{i,h}\right)_{h\in\mathcal{H}_i\wedge i\in\mathcal{N}\{0\}}.$$

This implies that also $BR^{\varepsilon}:\Sigma^{\varepsilon}\rightarrow\Sigma^{\varepsilon}$ is upper semi continues, compact and convex and, hence, has a fixed point $\hat{\sigma}^{\varepsilon}$. As already pointed out by Geanakoplos et al. (1989), the profile $\hat{\sigma}^{\varepsilon}$ constitutes an equilibrium of the constrained game Γ^{ε}.

Now, let ε^k be a sequence converging to 0 and $\hat{\sigma}^k$ the corresponding sequence of equilibrium assessments with $\hat{\sigma}^k$ being an equilibrium of Γ^{ε^k}. By the compactness of Σ, $\hat{\sigma}^k$ has an accumulation point σ^* and by the upper-semi-continuity of the local best-response-correspondents, $BR_h^{\varepsilon}(\sigma)$, $\sigma_{i,h}^*$ assigns positive probability only to those actions that are best responses to $(\sigma^*,\beta(\sigma^*),\omega)$ at h. Therefore $((\sigma^*,\omega),(\beta(\sigma^*),\omega))$ is a sequential equilibrium assessment. This concludes the proof. ■

In this section we have formally defined sequential psychological equilibria in the context of our class of games with psychological payoffs and moves of chance. Furthermore we have shown that every game with psychological incentives and continuous utility functions has at least one SPE. Using this solution concept we demonstrate in the following section the impact of procedural choices on the interaction of psychologically motivated agents in the principal-agent context.

4 The Impact of Appointment Procedures

In this section we formalize the principal-agent example already sketched in the introduction and analyze the potential impact of the appointment procedure on the performance of the typists using the sequential equilibrium concept defined in the previous section. Consider a two-stage game with a principal and two agents. In the first stage, the principal chooses the procedure by which the typist is chosen. The principal has three different strategies: He can directly appoint agent i (strategy d_i); he can directly appoint the other agent j (strategy d_j); he can choose a random appointment procedure (strategy r), where chance determines the typist with equal probabilities for both agents.

After the tasks are allocated, the appointed typist determines his effort e in the second stage. The minimum and the maximum effort levels are denoted by $e_{\min}$ and $e_{\max}$, respectively, with $e_{\min} \geq 0$. Associated with effort are effort costs $c(e) \geq 0$. In order to analyze the detrimental effect of the appointment procedure on the typist's effort choice, and in accordance with the empirical results shown in the next section, we assume that the typists will provide at least some effort voluntarily. This means that effort costs are minimized at some $\widetilde{e}$ with $e_{\min} < \widetilde{e} < e_{\max}$. The effort cost function is twice continuously differentiable and strictly convex. In order to avoid corner solutions we assume that $c'(e_{\min}) = -\infty$ and $c'(e_{\max}) = \infty$.

As already pointed out in the introduction, we are interested in the impact of the appointment procedure on the effort of the typist. Hence, we do not model any effort choice of the controller. The profit of the principal is given by

$$\pi_p = e - w_c - w_t,$$

where w_c and w_t denote the wages of the controller and the typist. Different to the example in section 2, assume the principal is not motivated by any fairness or reciprocity considerations. He is a pure profit maximizer.

As already mentioned, we assume the controller does not provide any effort. Hence, his direct 'material' payoff is w_c. This payoff is of course not relevant for any decision of the controller, since the controller has no decision to make. But as we will see it is important when the typist has to evaluate the decision made by the principal. Consequently, it influences the typist's effort choice.

The typist gets a lower wage than the controller, i.e. $w_t < w_c$. We assume that the wage difference is larger than the largest possible effort cost difference, i.e. $w_c - w_t > \max_e \left(c(e) - c(\widetilde{e})\right)$.

Agent i becomes a typist and has to choose his effort level whenever either the principal has chosen d_i, or the principal has chosen r and chance has appointed i as typist. Denote by $e_i(s)$, $s \in \{d_i, r\}$, the effort choice of agent i if the principal chooses s, and if chance appoints i in case of r. Disregarding any psychological payoff, agent $i's$ expected direct, 'material' payoff is given by

$$\pi_i(s, e_i(s)) = w_t - c(e_i(s))$$

with $s \in \{d_i, r\}$, and with nature appointing i in case of r.[9] But we assume that agents do not only care about their material payoffs, but are also motivated by belief-dependent psychological payoffs.

[9] During the whole analysis, we assume that the agents cannot quit the job. This is equivalent to assuming that the outside payoff is sufficiently small.

In order to get testable predictions, one has to further specify the type of belief-dependent psychological payoff. The specification we use captures e.g. reciprocity as modeled by Rabin (1992) and by Dufwenberg and Kirchsteiger (2004).[10]

To model these types of psychological payoffs, first- and second-order beliefs have to enter the utility functions. Denote by $\overline{\overline{e}}_i(s)$ the second-order belief of agent i about the principal's belief about agent i''s effort choice when the principal chooses s, $s \in \{d_i, r\}$, and when chance appoints i in case of r. Denote by $\overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(s))$ the belief of agent i about the material payoff the principal intended for i when i finds himself in a situation where he has to make an effort choice, i.e. when the principal chooses $s \in \{d_i, r\}$, and chance appoints i in case of r. Note that the equilibrium requires that the first-order belief of the agent about the principal's choice is equal to the actual choice of the principal. Therefore, in this simple game we do not have to distinguish notationally between the agent's first-order belief about the principal's choice and the actual choice of the principal. Note further, as specified in section 2, the agent does not hold the principal responsible for the choice of the moves of chance. Hence, $\overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(s))$ is given by

$$\overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(s)) = \begin{cases} \frac{1}{2}w_c + \frac{1}{2}(w_t - c(\overline{\overline{e}}_i(r)) & \text{if } s = r \\ w_t - c(\overline{\overline{e}}_i(d_i)) & \text{if } s = d_i. \end{cases}$$

Since $w_c - w_t > \max_e (c(e) - c(\widetilde{e}))$, $\overline{\overline{\pi}}_i(r, \overline{\overline{e}}_i(r)) > \overline{\overline{\pi}}_i(d_i, \overline{\overline{e}}_i(d_i)$ for any $\overline{\overline{e}}_i(r)$, $\overline{\overline{e}}_i(d_i)$.

Denote by $\overline{\pi}_p$ the principals material payoff agent i intend to give him whenever either the principal has chosen d_i, or the principal has chosen r and chance has appointed i as typist. $\overline{\pi}_p$ is given by

$$\overline{\pi}_p(s, e_i(s)) = \begin{cases} e_i(r) - w_c - w_t & \text{if } s = r \text{ and chance has chosen } i \\ e_i(d_i) - w_c - w_t & \text{if } s = d_i \end{cases}$$

In accordance with the literature we assume that agents' overall payoff is quasilinear in the own material payoff and in a psychological term depending on $\overline{\pi}_p$ and on $\overline{\overline{\pi}}_i$. If agent i finds himself in a situation where he is the typist, his utility is given by

$$u_i(s, e_i(s), \overline{\overline{e}}_i(s)) = w_t - c(e_i(s)) + v_i(\overline{\pi}_p(s, e_i(s)), \overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(s)))$$

for $s \in \{d_i, r\}$, and in case of r chance appoints i. We assume that v_i is twice differentiable, that the partial derivatives of v_i with respect to $\overline{\pi}_p$ and $\overline{\overline{\pi}}_i$ are finite, and that

$$\begin{aligned} \frac{\partial^2 v_i}{\partial \overline{\pi}_p \partial \overline{\overline{\pi}}_i} &> 0. \qquad (2) \\ \frac{\partial^2 v_i}{\partial \overline{\pi}_p^2} &\leq 0 \end{aligned}$$

[10]Note, on could also assume that agents are motivated by other emotions like guilt á la Battigalli and Dufwenberg (2007) or dissapointment á la Ruffle (1999). However, it can be shown that for such emotions multiple SPE exist. Hence, these models would not provide us with testable predictions.

This assumption captures the essence of fairness and reciprocity considerations. It states that the marginal utility derived from the material payoff the agent intends to give to the principal increases in the material payoff the agent thinks that the principal wants to give to him. Furthermore, the marginal utility derived from the payoff intended to be given to the principal does not increase in the principal's intended payoff. As said above and as can be easily checked, the reciprocity theories of Rabin (1992) and Dufwenberg and Kirchsteiger (2004) fulfill this assumption.

Using this type of psychological motivation, we get the following result:

Proposition 2 *In any sequential psychological equilibrium, $e(r) > e(d_i)$.*

Proof: When agent i has to make an effort choice, he maximizes his utility for given s and given $\overline{\overline{\pi}}_i$. Formally, for $s \in \{d_i, r\}$ the maximization problems read

$$\max_{e_i(s)\in[e_{\min},e_{\max}]} w_t - c(e_i(s)) + v_i(\overline{\pi}_p(s, e_i(s), \overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(s)))$$

Since $\frac{\partial \overline{\pi}_p(s,e_i(d_i))}{\partial e_i(d_i)} = \frac{\partial \overline{\pi}_p(s,e_i(r))}{\partial e_i(r)} = 1$ the first order conditions are given by

$$-c'(e_i(d_i)) + \frac{\partial v_i(\overline{\pi}_p(s, e_i(d_i), \overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(d_i)))}{\partial \overline{\pi}_p} = 0 \quad (3)$$

and

$$-c'(e_i(r)) + \frac{\partial v_i(\overline{\pi}_p(s, e_i(r), \overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(r)))}{\partial \overline{\pi}_p} = 0. \quad (4)$$

Since $c(e_{\max}) = -c(e_{\min}) = \infty$, the maximization problem has only interior solutions for any $s \in \{d_i, r\}$, and the optimal effort choices have to fulfill these FOC's.

Now assume that contrary to the proposition, there exists an equilibrium with $e(r) \leq e(d_i)$. Because of strict convexity of the effort cost function, this implies that $-c'(e_i(r)) \geq -c'(e_i(d_i))$. Because of the first order conditions, we get

$$\frac{\partial v_i(\overline{\pi}_p(s, e_i(r), \overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(r)))}{\partial \overline{\pi}_p} \leq \frac{\partial v_i(\overline{\pi}_p(s, e_i(d_i), \overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(d_i)))}{\partial \overline{\pi}_p} \quad (5)$$

Recall that $\overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(r)) > \overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(d_i))$ for any $\overline{\overline{e}}_i(r)$, $\overline{\overline{e}}_i(d_i)$. Together with 2), 5) implies

$$\overline{\pi}_p(s, e_i(d_i)) < \overline{\pi}_p(s, e_i(r)).$$

But this condition only holds for $e_i(d_i) < e_i(r)$ - a contradiction.■

We have shown that whenever the agents have belief-dependent preferences as defined above and thus care about the principal's choice of procedure, the typist picked by a random mechanism should work harder than the one directly appointed. The intuition behind that result is straightforward: If picked by the random mechanism and not on purpose, the agent attributes 'better' intentions to the principal's choice (i.e. $\overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(r)) > \overline{\overline{\pi}}_i(s, \overline{\overline{e}}_i(d_i))$, and is therefore willing to work harder. He does not hold the principal responsible for being typist, it is just bad luck.

5 The field experiment

5.1 Setup

To test our result in proposition 2, i.e. the impact of the principal's procedural choice on the effort choice of the disadvantaged agent, we conducted a field experiment at the University of Namur. We hired research assistants (RAs) for an ongoing research project that involves constructing a large dataset on the evolution of family structures in XIX-XXth century Russia and Kazakhstan. Half of the research assistants (the 'typists') had to type numerical data into a Microsoft Excel worksheet from photocopies of old statistical publications of the Russian Empire. The others (the 'controllers') had to check whether the data typed in was correct. All RAs were employed for two hours. The typists received a flat hourly wage of €10, whereas the controllers received a flat hourly wage of €15. There is no obvious difference in terms of intrinsic (dis)utility of labor for both jobs. If anything, the controllers's task seems to be less unpleasant. Taking the wage difference into account, the typist's job is for sure less attractive than that of the controller.

As also laid out in section 4, to test for procedural concerns, we concentrate in our experiment on the performance of the typists who are disadvantaged compared to the controllers. If procedural concerns play no role, the performance of typists (in terms of the amount of data typed in and of typos made) should be independent of the way they are appointed. We used two different mechanisms to appoint the typists and the controllers: a direct and a random appointment mechanism. In the treatment with the 'direct appointment procedure', we directly appointed RAs to typist and controller roles (without giving any justification for the appointment to a given role). In the treatment with the 'random appointment procedure', each RA drew a card from a bowl to determine her/his role. There were equally many typist and controller cards in the bowl giving each RA an equal chance to become typist or controller.

Under both procedures, half of RAs were appointed as typists and the other half as controllers. All RAs were made aware of the wage difference, as well as of the fact that one-half of them were controllers and the other half were typists (see the instructions in the Appendix). Note, however, that none of the participants was made aware of the fact that this was not only a real RAs' job but also an experiment.

We hired freshmen and sophomore undergraduate students of the University of Namur studying in six different faculties (Economics, Law, Science, Computer Science, Philosophy, and Medicine) as RAs. For each RA, we know which faculty and year s/he studies in, whether the student is foreign-born, and his or her gender. Note that, as some registered students did not show up on the date of experiment, the numbers of RAs subject to the two appointment procedures differs slightly.

5.2 Summary statistics

Table 1 presents the distribution of typists across treatments and gender. Overall, there were 43 typists: 24 appointed directly and 19 appointed through the random appointment procedure. 25 subjects were men and 19 were women. The number of male subjects in two treatments was roughly equal (13 under direct appointment and 12 under random appointment), whereas for women there was a slight over-representation in the direct-appointment treatment (11 versus 7 under random appointment).

[Insert Table 1 here]

Table 2 presents the summary statistics for our three measures of performance. On average, in two hours of work, a typist encoded 3675 cells in the Excel worksheet. However, there is substantial variation in the number of cells typed in: standard deviation is 1031 cells, with the minimum equal to barely over 2000 cells and the maximum over 6800 cells.

[Insert Table 2 here]

We also measured the number of cells typed in incorrectly ('typos'), using the typos detected by controllers, which were also cross-checked by other research assistants. On average, a typist made 6.74 typos in two hours. Again, the performance varied substantially: standard deviation was 5.42 typos, with some typists making 0 mistakes, while some making as many as 20 typos.

Clearly, a typist typing in more data was also likely to make more typos. To account for this, we also measure the error rate. This is a common measure of performance in statistical quality control (Montgomery 2008). On average, a typist typed in 0.19% of cells incorrectly. This error rate might look small, but actually a large part of the cells had to be filled by zeros. In these cases no effort was required to avoid an error. So the error rates for the "effort demanding" cells were actually much larger. Furthermore, there was still large variation in the error rate: standard deviation is 0.16%, with the minimum error rate equal to 0 and the maximum error rate of 0.6%.

5.3 Experimental results

Tables 3-5 present our experimental results.

[Insert Tables 3-5 here]

As one can see from Table 3, an average typist inserted 3470 cells under the 'direct appointment' procedure and an average typist inserted 3934 cells under 'random appointment' procedure. Men typed in 3892 cells on average, while women typed in 3375 cells on average.

However, the results across treatments were strikingly different for men and women. Women inserted on average 3444 cells under 'direct appointment' and a somewhat lower number of cells - 3267 - under 'random appointment'. Men, on the contrary, inserted substantially more cells under 'random appointment' (4324, on average) than under 'direct appointment' (3494, on average). Thus, in terms of the number of cells typed in, there was an important effect of procedures on the performance of male typists, with 'random appointment' procedure inducing higher performance, while for female typists the effect was weak and in the opposite direction.

Table 4 presents the results on the number of typos made. Here, the results are similar for both men and women. Male (female) typists appointed using the random appointment procedure made fewer typos (4.8 and 4.6, respectively) as compared to those chosen using 'direct appointment' (9.3 and 7.2, respectively). The effect when both genders are considered is 4.7 typos on average under 'random appointment' versus 8.3 typos under 'direct appointment'.

Finally, the quality control results are presented in Table 5. They are similar to those on the number of typos. On average, typists hired under the 'random appointment' procedure worked with an error rate of 0.12%, while those hired under 'direct appointment' had an error rate twice as high, namely 0.24%. For men, the corresponding error rates were 0.11% versus 0.26%, and for women they were 0.14% versus 0.23%.

5.4 Regression results

We now proceed to more rigorous statistical analysis. We estimate the following econometric model:

$$y_i = \alpha + \beta \mathbf{I}_i(r = 1) + \gamma \mathbf{X}_i + \varepsilon_i, \tag{6}$$

where y_i is the measure of individual performance of typist i, $\mathbf{I}_i(r = 1)$ is the indicator variable that takes value 1 if the typist is chosen under the 'random appointment' procedure and 0 if she/he is appointed under 'direct appointment", $\mathbf{X}_i$ is the vector of individual characteristics that might capture a part of the variation in the performance, and ε_i is the error term that we assumed to be normally distributed with zero mean and a constant variance.

Our theoretical model predicts that the coefficient β is positive and significantly different from zero in the statistical sense. In other words, finding a positive and statistically significant β would mean that once we hold other observable individual characteristics constant, individuals in the 'random appointment' procedure exhibit higher performance than in the 'direct appointment' procedure.

Moreover, given that the above results suggest that typists of different gender might respond differently to the same procedure, we will also estimate the following altered model:

$$y_i = \alpha + \beta \mathbf{I}_i(r = 1) + \delta \mathbf{G}_i(f = 1) + \mu[\mathbf{I}_i(r = 1) * \mathbf{G}_i(f = 1)] + \gamma \mathbf{X}_i + \varepsilon_i, \tag{7}$$

where $\mathbf{G}_i(f = 1)$ is the indicator variable for gender (which takes value 1 is the typist is a woman and 0 otherwise). The differential-by-gender response to the 'random appointment' procedure is thus captured by the coefficient μ. Thus, finding a positive (negative) and statistically significant μ would mean that once we hold all other observable individual characteristics constant, women in the 'random appointment' procedure exhibit higher (lower) performance than in the 'direct appointment' procedure.

Table 6 presents the estimation results of models (6) and (7).

[Insert Table 6 here]

Columns (1)-(3) show results with the number of cells inserted as the measure of individual performance. Column (1) reports the results of the estimation with only the treatment as a regressor. On average, typists appointed using 'random appointment' encode 464 cells more as compared to those chosen by 'direct appointment', but this difference is not statistically significant. The results become much more clear-cut when we add the gender of the typist and allow for the differential-by-gender response to the treatment. Column (2) reports the results of the estimation of the amended model (7), without additional controls. A typist hired under 'random appointment' inserts 830 cells more as compared to those hired under 'direct appointment', and this difference is significant at 5%. However, if the typist is a woman, under 'random appointment' she inserts 226 cells less ($830 - 50 - 1006 = -226$) than under 'direct appointment'. Clearly, there is substantial variation in individual performance, a part of which is captured when we add the additional controls that might be correlated with unobservable skills: faculty and freshmen dummies (people sort into faculties and there is probably some acquisition of skill over time) and the foreign-born dummy (a student whose mother tongue is not French might take more time to understand the instructions, which might reduce his/her performance). Column (3) presents the results of the estimation of the model (7) with these additional regressors. The coefficients β and μ both increase in absolute value and are more precisely estimated (both are significant at 1%). Moreover, the adjusted-R^2 of the model is the highest among the three estimations. Thus, a male typist under 'random appointment' encodes 1327 cells more than under the 'direct appointment'. This is a quantitatively large effect, about 1.3 times the standard deviation. The effect of the 'random appointment' for a women is, however, small and negative: $1327 + 341 - 1970 = -302$.

Columns (4)-(6) and (7)-(9) present the results of the estimation with the number of typos and the error rate as the measure of performance. In both cases, the effect of the appointment procedure is clear and similar for both male and female typists (the coefficient μ is not significantly different from zero in any estimation). The results are similar across all the estimations. Using the model with the best fit (as measured by the adjusted-R^2), we can state that a typist appointed using the 'random appointment' procedure makes 3.6 typos less (and has an error rate of 0.12% lower) than if the 'direct appointment' procedure is used. This effect is quantitively large: it equals 2/3 of the standard deviation in the case of typos and 3/4 in the case of the error rate.

Summarizing, in line with proposition 2 our experimental results suggest that the appointment procedure has significant and large effects on individual performance. However, the form of the effect differs for men and women. If hired under 'random appointment', men exhibit higher performance by all counts: both in terms of quantity and quality of output. Instead, women - if hired under 'random appointment' - increase the quality of their output but not the quantity.

These findings provide clear support for our theoretical hypotheses. They indicate that in the field-experimental setting procedures matter for individual performance: research assistants working in a natural environment and unaware of participating in an experiment exhibit substantial differences in terms of quality and quantity of effort that they exert, depending on the way we - the experimenters - have chosen to assign them the worse task.

Moreover, our findings in terms of gender differentials in effort (in quantity, but not quality of effort) complements the existing literature on the gender differences in social preferences. Croson and Gneezy (2009, section 3) argue, based on their interpretation of a large body of the experimental literature that women seem to be more sensitive to the experimental context. Our results qualify this argument: in the principal-agent field-experimental setting, both women and men seem to exert less effort if the procedure chosen by the principal is considered less fair; however, women carry out this reduction of effort in a subtler way than men. Women do not reduce the quantity of output, but reduce its quality, while men reduce their effort more explicitly, i.e. both the quantity and quality.

6 Conclusion

In this paper we have shown how decision-making procedures can be integrated into the framework of dynamic psychological games by Battigalli and Dufwenberg (2009) by allowing for moves of chance. After proving the existence of sequential psychological equilibria within this extended framework, we investigated the impact of procedural concerns in the context of a specific principal-agent relation. Our model predicted that whenever agents have belief-dependent preferences, the principal's choice of appointment mechanism is crucial for the elicited effort. We then tested this hypothesis in a field experiment, and its results confirm our predictions. Moreover, we find interesting gender differences in the form of the reaction to different procedures.

While our paper provides a general framework for procedural concerns, the application to the principal-agent context as well as its experimental test concentrate on specific belief-dependent preferences. These specific belief-dependent preferences (which encompass, for instance, reciprocity) are used to come up with testable predictions concerning the impact of procedural choices on the effort provision of agents in our principal-agent setting. However, as we have discussed, procedural concerns are not confined to these specific belief-dependent motivations but also arise under other psychological incentives (guilt, disappointment, etc.). The specific analyses of the impact of procedural choices on the interaction of agents with these other types of belief-dependent motivations is left for future research.

7 References

1. Anderson, R. and Otto, A. (2003), *Perceptions of fairness in the justice system: A cross-cultural comparison*, Social Behavior and Personality, 31, 557-564.

2. Aumann, R. and Brandenburger, A. (1995), *Epistemic conditions for Nash equilibrium*, Econometrica, 63, 1161-1180.

3. Battigalli, P. and Siniscalchi, M. (1999), *Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games*, Journal of Economic Theory, 88, 188-230.

4. Battigalli, P. and Dufwenberg, M. (2009), *Dynamic Psychological Games*, Journal of Economic Theory, 144, 1-35.

5. Battigalli, P. and Dufwenberg, M. (2006), *Guilt in Games*, American Economic Review, Papers and Proceedings, 97, 170–176.

6. Bies, R. and Tyler, T. (1993), *The 'litigation mentality' in organizations: A test of alternative psychological explanations*, Organization Science, 4, 352-366.

7. Blader, S. and Tyler, T. (2003), *A four-component model of procedural justice: Defining the meaning of a 'fair' process*, Personality and Social Psychology Bulletin, 29, 747-758.

8. Blount, S. (1995), *When Social Outcomes Aren't Fair: The Effect of Casual Attributions on Preferences*, Organizational Behavior and Human Decision Processes, 63, 131-144.

9. Bolton, G., Brandts, J. and Ockenfels, A. (2005), *Fair Procedures: Evidence from Games Involving Lotteries*, Economic Journal, 115(506), 1054-1076.

10. Brandenburger and Dekel (1993), *Hierarchies of Beliefs and Common Knowledge*, Journal of Economic Theory, Elsevier, 59(1), 189-198.

11. Brandts, J., Gueth, W. and Stiehler, A. (2006), *I want you!: An experiment studying the selection effect when assigning distributive power*, Journal of Labor Economics, 13, 1-17.

12. Charness, G. and Rabin, M. (2002), *Understanding Social Preferences With Simple Tests*, Quarterly Journal of Economics, 117(3), 817-869.

13. Charness, G. (2004), *Attribution and Reciprocity in an Experimental Labor Market*, Journal of Labor Economics, 22(3), 553-584.

14. Charness, G. and Dufwenberg, M. (2006), *Promises and Partnership*, Econometrica, 74(6), 1579-1601.

15. Charness, G. and Levine, D. (2007), *Intention and Stochastic Outcomes: An Experimental Study*, Economic Journal, 117(522), 1051-1072.

16. Collie, T., Bradley, G. and Sparks, B. (2002), *Fair process revisited: Differential effects of interactional and procedural justice in the presence of social comparison information*, Journal of Experimental Social Psychology, 38, 545-555.

17. Croson, R., and Gneezy, U. (2009), *Gender differences in preferences*, Journal of Economic Literature 47(2): 448-74.

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

19. Falk, A., Fehr, E. and Fischbacher, U. (2000), *Testing Theories of Fairness - Intentions Matter*, Working Paper No. 63, University of Zurich, forthcoming in: Games and Economic Behavior.

20. Fehr, E. and Gaechter, S. (2000), *Fairness and Retaliation: The Economics of Reciprocity*, Journal of Economic Perspectives, 14(3), 159-181.

21. Geanakoplos, J., Pearce, D. and Stacchetti, E. (1989) *Psychological Games and Sequential Rationality*, Games and Economic Behavior, 1, 60-79.

22. Konovsky, M. (2000), *Understanding Procedural Justice and Its Impact on Business Organizations*, Journal of Management, 26(3), 489-511.

23. Lemons, M. and Jones, C. (2001), *Procedural justice in promotion decisions: using perceptions of fairness to build employee commitment*, Journal of Managerial Psychology, 16(4), 268-281.

24. Lind, E. and Tyler, T. (1988), *The social psychology of procedural justice*, New York, NY, US: Plenum Press.

25. Lind, E., Greenberg, J., Scott, K. and Welchans, T. (2000), *The winding road from employee to complainant: Situational and psychological determinants of wrongful termination claims*, Administrative Science Quarterly, 45, 557-590.

26. Montgomery, D. (2008), *Introduction to Statistical Quality Control*, 6th ed. New York: Wiley.

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

28. Roberts, K. and Markel, K. (2001), *Claiming in the name of fairness: Organizational justice and the decision to file for workplace injury compensation*, Journal of Occupational Health Psychology, 6, 332-347.

29. Ruffle, B.J. (1999), Gift giving with emotions, Journal of Economic Behavior Organization, 39, 399–420.

30. Sebald, A. (2009), *Attribution and Reciprocity*, Mimeo, University of Copenhagen, Copenhagen.

31. Thibaut J. and Walker, L. (1975), *Procedural Justice*, Erlbaum, Hillsdale, NJ.

8 Appendix

The following information was read out to subjects at the beginning of each session. The contents were identical in both treatments, except the section marked in italics.

Job description and payment details

We are constructing a dataset on the socio-economic characteristics of extended families and their production and consumption decisions, for a project on the evolution of family structure and collective action in traditional societies. The raw data that we have (that comes from an agricultural census of the Russian Empire of the beginning of the 20th century) exists only in the paper version, and not in electronic format. This means that it is necessary to copy it from the paper version into an Excel worksheet.

[Detailed instructions on how to copy the data from the paper version into the worksheet]

To make sure that the data is inserted correctly, all the files will be crosschecked. This means that there are two different tasks. TYPISTS insert data into Excel worksheets. CONTROLLERS verify the inserted data and correct it wherever necessary.

The hourly wage is 15€ for a controller and 10€ for a typist. In total, you are going to work for 2 hours; thus, a typist will receive 20€ and a controller 30€ at the end of the work.

[TREATMENT 1] Given the lack of time, we cannot verify which of you are better qualified to work as a typist or as a controller. We thus have decided that you are going to work as a TYPIST.

[TREATMENT 2] Given that we do not know which of you are better qualified to work as a typist or as a controller, the tasks are allocated in a random fashion. Each of you had to draw a card from the bowl. If you have picked a card with the word "TYPIST", you are going to work as a typist. If you have picked a card with the word "CONTROLLER", you are going to work as a controller.

In order to avoid losing the data inserted, please make sure that you save your Excel file regularly.

Do you have any questions?

Table 1. Number of observations

	Male	Female	Both
Direct appointment	13	11	24
Random appointment	12	7	19
Both treatments	25	18	43

Table 2. Summary statistics

	Mean	Std. Dev.	Min	Max
Number of cells encoded	3675	1031	2010	6825
Number of typos made	6.74	5.42	0	20
Error rate, in %	0.19	0.16	0	0.6

Table 3. Average number of cells encoded

	Male	Female	Both
Direct appointment	3494	3444	3470
Random appointment	4324	3267	3934
Both treatments	3892	3375	3675

Table 4. Average number of typos made

	Male	Female	Both
Direct appointment	9.3	7.2	8.3
Random appointment	4.8	4.6	4.7
Both treatments	7.2	6.2	6.7

Table 5. Average error rate, in %

	Male	Female	Both
Direct appointment	0.26	0.23	0.24
Random appointment	0.11	0.14	0.12
Both treatments	0.18	0.19	0.19

Table 6. Multiple Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	Number of cells encoded	Number of cells encoded	Number of cells encoded	Number of typos made	Number of typos made	Number of typos made	Error rate, in %	Error rate, in %	Error rate, in %
Constant	3470	3494	2432	8.33	9.31	8.34	0.24	0.26	0.26
	(16.73)***	(12.86)***	(3.59)***	(7.90)***	(6.41)***	(2.21)**	(8.18)***	(6.24)***	(2.35)**
Treatment = Random appointment	464	830	1327	-3.60	-4.47	-3.66	-0.12	-0.15	-0.15
	(1.49)	(2.12)**	(2.95)***	(2.27)**	(2.14)**	(1.46)	(2.75)***	(2.50)**	(2.06)**
Female		-50	341		-2.21	-0.71		-0.03	-0.01
		(0.12)	(0.78)		(0.99)	(0.29)		(0.50)	(0.15)
Female*Random		-1006	-1970		1.86	-0.35		0.06	0.05
		(1.64)	(2.73)***		(0.57)	(0.09)		(0.67)	(0.39)
Controls	No	No	Yes	No	No	Yes	No	No	Yes
Observations	43	43	43	43	43	43	43	43	43
Adjusted R-squared	0,03	0,10	0,15	0,09	0,07	0,05	0,14	0,10	0,03

Absolute value of t-statistics in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Treatment: omitted category = "Direct appointment"
Controls: Faculty, freshman, and foreign-born dummies