Moral Hazard and Other-Regarding Preferences*

Hideshi Itoh†

First Version: June 30, 2003

This Version: September 12, 2003

*I am grateful to Munetomo Ando, Shingo Ishiguro, Shinsuke Kambe, Michihiro Kandori, Hitoshi Matsushima, Masahiro Okuno-Fujiwara, Pedro Rey Biel, Takashi Shimizu, Katsuya Takii, an anonymous referee, and the participants at Contract Theory Workshop (CTW), Contract Theory Workshop East (CTWE), and University of Tokyo Microeconomics Workshop for helpful comments. The research for this paper was supported by Grant-in-Aid for Scientific Research (KAKENHI 13630005) from JSPS and MEXT of the Japanese Government.

†Graduate School of Commerce and Management, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan.

E-mail: h.itoh@srv.cc.hit-u.ac.jp. Url: http://obata.misc.hit-u.ac.jp/˜itoh/index-e.html

Abstract

The paper aims at obtaining new theoretical insights into organizational behavior by combining the standard moral hazard models of principal-agent relationships with theories of other-regarding (social or interdependent) preferences, in particular, inequity aversion theory. In the benchmark model, the principal and the agent are both risk neutral, while the agent is wealth constrained and hence the basic tradeoff between incentives and rent extraction arises. I show that other-regarding preferences interact with incentives in nontrivial ways. In particular, the principal is in general worse off as the agent cares more about the well-being of the principal. When there are multiple symmetric agents who care about each other's well-being, the principal can optimally exploit their other-regarding nature by designing an appropriate interdependent contract such as a "fair" team contract or a relative performance contract that creates inequality when their performance outcomes are different. The optimal contract depends on the nature of the agents' other-regarding preferences. The approach taken in this paper can shed light on issues on endogenous preferences within organizations, as suggested by sociologists and organizational economists.

JEL Classification Numbers: D82, D63, M52, M54.

Keywords: Principal, agent, incentives, other-regarding preferences, inequity aversion, status seeking, team, relative performance, behavioral contract theory.

1 Introduction

It is standard in economic analysis to assume that people maximize their wealth and other personal material consumption. This self-interest hypothesis has proved to be correct in many situations and highly useful in the analysis of diverse problems. However, people, certainly including economists, have recognized that *not all* people are *exclusively* motivated by their self-interest. We often care about the well-being of others, and find ourselves behaving altruistically, worrying about unfair distribution of wealth, reciprocating kind or unkind behavior of others, and so on. These sorts of *other-regarding* behavior, in addition to selfish behavior, are the concern of this paper.

The recent data from experiments on ultimatum games, gift exchange games, public goods games, trust games, and so on, demonstrate that people in fact deviate from self-interest in systematic ways.[1] "Among experimentalists—and others paying attention to the evidence—the debate over whether there are systematic, non-negligible departures from self-interest is over (Rabin, 2002, p.666)." In particular, the data from a large number of ultimatum game experiments "falsify the assumption that players maximize their own payoffs as clearly as experimental data can. Every methodological explanation you can think of (such as low stakes) has been carefully tested and cannot fully explain the results (Camerer, 2003, p.43)."

Based on the experimental evidence, researchers in behavioral economics and behavioral game theory have recently developed new theoretical models that deviate from the exclusive self-interest hypothesis by incorporating other-regarding features such as fairness, equity, and reciprocity, in the framework of the standard utility maximization. Behavioral economics is "an approach to economics which uses psychological regularity to suggest ways to weaken rationality assumptions and extend theory (Camerer, 2003, p.3)." Some of the new models capture other-regarding behavior, and succeed in explaining various experimental results.

The purpose of this paper is to go one step further. I apply those theories of other-regarding preferences that can explain many experimental results to the standard models of principal-agent relationships with moral hazard, in order to generate new theoretical insights. The traditional literature in contract theory is also based on the self-interest hypoth-

[1] See Camerer (2003) and Fehr and Schmidt (2003) for surveys.

esis. The agent attempts to maximize a function that is increasing in his wealth and other private benefits and is decreasing in the private cost of his action. The principal aims at maximizing the benefit generated by the agent minus her payments to him.[2]

However, relaxing the self-interest hypothesis may be particularly important for contract theory. First, the theme of contract theory is designing appropriate incentives, and how people care about others' well-being as well as their own is crucial for incentive design. The optimal contract for the self-interested agent may be very different in nature from that for the other-regarding agent.

Second, contract theory offers a major analytical framework for organizational economics (Milgrom and Roberts, 1992) and personnel economics and human resource management (Lazear, 1995; Baron and Kreps, 1999). Even in these applied fields, the dominant view is that "many institutions and business practices are designed as if people were entirely motivated by narrow, selfish concerns and were quite clever and largely unprincipled in their pursuit of their goals (Milgrom and Roberts, 1992, p.42)." However, organizational researchers in other disciplines have criticized such a view over the years (Perrow, 1986; Pfeffer, 1997). For example, Perrow (1986) argues that "there is no innate tendency to either self- or other-regarding behavior in people; either can be evoked depending on the structure. Agency theorists examine the structures favored by capitalism and bureaucracy and find much self-regarding behavior; they then assume that this is human nature. They neglect the enormous amount of neutral and other-regarding behavior that exists (and must, for organizations even to function) and the structures that might increase it. What they take for granted, should be taken as a problem (Perrow, 1986, pp.234–235)."

My optimistic standpoint is that developing "behavioral contract theory" that uses experimental/field evidence and psychological intuition to generalize the standard assumptions could contribute to encouraging more productive interactions among the students of organizations from various disciplines. As a start, this paper analyzes how incentives are affected

[2]Even in the traditional contract theoretic models, other-regarding behavior sometimes plays an implicit but important role. In the collusion literature (Holmstrom and Milgrom, 1990; Itoh, 1993; Tirole, 1992), the enforceable side-contracts assumption may reflect other-regarding behavior between agents implicitly. Some of the important results from the multitask analysis (Holmstrom and Milgrom, 1991) hinge on the assumption that the agent chooses some positive amount of effort in the absence of incentives. This feature may be explained by task-specific intrinsic motivation, or the agent's caring about the well-being of the principal.

by other-regarding behavior in the stylized principal-agent relationships with moral hazard, and what kinds of preferences are desirable for the principal.

Third, very few experiments on contract choice by the principal who cannot observe the agent's action have so far been conducted.[3] Providing theoretical results on how incentives and other-regarding behavior interact in principal-agent relationships will help promote experimental tests of contract theory in future research.

To study the effects of other-regarding preferences on moral hazard, I mainly use the model of *inequity aversion* developed by Fehr and Schmidt (1999).[4] Their model is an example of the distributional approach to other-regarding preferences in which only the final monetary distribution among players matters. In particular, a player feels guilty when his material payoff is above others' payoffs, while he feels envious when his material payoff is below others'. In other words, he dislikes either being ahead or being behind. The model is simple but can explain various experimental results and capture fair-minded and reciprocal behavior very well.[5] Furthermore, to cover broader cases of other-regarding preferences, I follow Neilson and Stowe (2003) to extend Fehr and Schmidt (1999) by allowing the player to be *competitive* or *status-seeking* in the sense that he dislikes being behind, but loves to be ahead.

The benchmark moral hazard model is a standard one. A principal hires an agent for a project. The project either succeeds or fails, and the probability distribution depends on the agent's action. I assume that both principal and agent are risk neutral while the agent is wealth constrained and hence the contract must satisfy the limited liability constraints. To induce the agent to choose the more productive action, the principal must offer high-powered incentives through a higher pay upon the success of the project. However, higher-powered incentives are more costly to the principal because she has to leave more rents to the agent.[6]

[3]Exceptions include Fehr et al. (2001), Güth et al. (1998), and Keser and Willinger (2000), all of which study trade between one principal and one agent. Nalbantian and Schotter (1997) present an interesting experimental examination of a variety of group incentives, although incentive plans are exogenously chosen by the experimenters.

[4]Bolton and Ockenfels (2000) present an alternative formulation of inequity aversion.

[5]Other kinds of distributional preferences as well as an alternative approach to other-regarding preferences are summarized in Section 2.

[6]In other words, the ex ante participation constraint does not bind. If I did not impose the limited liability constraints, the participation constraint would usually bind, and hence the effects of other-regarding preferences

I first assume that the agent compares his income with that of the principal. This setting corresponds to ultimatum games, gift exchange games, and trust games the experimental results of which are known to be consistent with the hypothesis that the player on the role of the receiver (agent) cares about the well-being of the player on the role of the proposer (principal), although the proposer is not allowed to offer complicated contracts. And Bewley (1999), who attempts to learn why wages and salaries seldom fall during recessions through interviews covering more than 300 people, finds some evidence that workers view pay cut as unfair in comparison with the performance of the company.

Other-regarding preferences interact with moral hazard in nontrivial ways. First, fairness concern does not resolve moral hazard: If the incentive compatibility constraint binds for the self-interested agent, it also binds for the other-regarding agent. Furthermore, in the "standard" case where the agent's income is below that of the principal, the principal is typically worse off as the agent is more other-regarding. The logic is simple: Since the agent envies inequity when the project succeeds, the principal must pay more upon the success of the project in order to satisfy the incentive compatibility constraint as the agent is more inequity averse.

Although at optimum the agent's income is lower than that of the principal under a reasonable assumption, we could instead assume the nonstandard situation that the agent receives more income than the principal under the successful project. Even in this case, the principal does not benefit from inequity aversion, because the inequity averse agent would be altruistic towards the principal and would want to reduce the probability of success (that is, the probability of being ahead). However, I show that if the agent has status-seeking preferences so that he enjoys being ahead, the principal is better off as the agent is more other-regarding (more competitive): As the agent is more status-seeking, he is more intrinsically motivated and hence the principal can save monetary incentives more.

The results obtained are mostly robust to changing specification of other-regarding preferences such that the agent compares his income net of costs of action with the principal's wealth, or that the principal as well as the agent is other-regarding.

In my view, it is more important to analyze the effects of other-regarding preferences in a multi-agent situation. A large body of sociological literature argues that people are most

on incentives, which are the focus of the paper, would be somehow undermined.

likely to compare themselves to others who are similar in terms of personal characteristics (such as age, gender, and education), and situations within organization (such as job titles, departments, and entry cohorts): While an agent may care about the principal's well-being, he is likely to care even more about others at the agent's position. For example, pay attached to job titles and pay differences across job titles are more likely to be perceived as acceptable by employees than pay differences within a single job title (Baron, 1988) That is, if multiple candidates for reference actors exist, then "lateral" comparison is more likely to dominate. I thus extend the model to a multi-agent setting, assuming that each agent cares about what he and the other agents are paid while the principal is not his reference actor.

I show that although I do not assume any technological or stochastic interdependence, the principal can optimally exploit the agents' other-regarding nature by designing an appropriate interdependent contract: Other-regarding preferences lead to the possibility that a team contract (each agent is paid more if the other's project succeeds than fails) or a relative performance contract (he is paid more if the other project fails than succeeds) becomes optimal. I obtain conditions for each a team contract and a relative performance contract to be optimal, and examine how the principal's payoff changes with the extent to which the agents are other-regarding. The "extreme" team contract under which each agent is paid a positive amount only when both projects succeed has the important feature that no agent feels guilty or envious, and hence the principal is indifferent concerning how other-regarding the agents are. However, the "extreme" relative performance contract (under which each agent is paid a positive amount only if his project succeeds and the other's fails) becomes optimal for status-seeking agents, or inequity averse agents who do not feel much guilty when being ahead. The reason such a contract becomes optimal is in contrast to that for the team contract. The relative performance contract creates large inequality in payments when the performance outcomes differ across agents, which feature provides strong incentives. This incentive effect is stronger as the agents are more other-regarding, and hence the principal benefits from more competitive agents.

The optimal contract for self-interested agents changes drastically when a small degree of other-regarding preferences is introduced. Under the technological assumptions of the model, if the agents are self-interested, there is an optimal independent contract in which the payment scheme for each agent depends only on the outcome of his project. However, when the agents become other-regarding, however small the changes are, no independent contract is optimal any longer and the optimal contract is generically unique. This result warns us against the use of independent contracts in the analysis of agency relationships with purely self-interested agents.

The analysis of the multi-agent setting seems to suggest that status-seeking agents be desirable, partly because of a lack of productive interaction among agents in the model. However, this result depends crucially on the specification that the agents compare only what they are paid. I show that if instead the agents compare their actions as well as the payments, offering an appropriate extreme team contract to sufficiently inequity averse agents is most desirable for the principal, because they are so intrinsically motivated to avoid being ahead that the principal can induce the agents to choose appropriate actions without leaving any rent to them.

Some recent literature has started to incorporate inequity aversion into principal-agent models. Englmaier and Wambach (2002) analyze the relationship between a principal and an agent. The agent is inequity averse and the principal constitutes a reference actor for the agent. They do not require contracts of satisfying the limited liability constraints, and hence in their analysis the participation constraint binds and plays a crucial role. Grund and Sliwka (2002) study rank-order tournaments among inequity averse agents, and show that inequity averse agents exert higher effort than purely self-interested agents for a given contract, while the first-best effort is not implementable when contracts are endogenous. Neilson and Stowe (2003) analyze the optimal linear contract for other-regarding and risk averse multiple agents, who are either inequity averse or status-seeking. The limited liability constraints are not imposed. More importantly, the principal in their model is only allowed to offer independent contracts.

In a paper written independently and concurrently with this one, Rey Biel (2003) also studies a relationship of a principal with two agents who are inequity averse between them, and shows that the optimal contract is either a team contract or a relation performance contract, although he does not use these terms. In contrast to my model, however, output is deterministic and perfectly informs of the actions chosen by the agents in his model. Furthermore, he exogenously assumes that the participation constraint does not bind. Instead, he allows agents to be asymmetric in terms of output and cost, and studies the implementation of independent production in which only one of the agents produces as well as of joint production.

The rest of the paper is organized as follows. In Section 2, I briefly summarize the recent theories of other-regarding preferences. In Section 3, I study the single agent case. Section 4 is the extension to a multi-agent setting. In Section 5 I discuss implications for choice of preferences within organizations. Section 6 is concluding remarks.

2 Theories of Other-Regarding Preferences

In this section, I briefly summarize the recent development in modeling other-regarding players by relaxing the pure self-interest assumption.[7] The purpose of each theory "is *not* to explain every different finding by adjusting the utility function just so; the goal is to find parsimonious utility functions, supported by psychological intuition, that are general enough to explain many phenomena in one fell swoop, and also make new predictions (Camerer, 2003, p.101)."

I call people *purely self-interested* if they care about their own material payoff, which I assume is monetary although it can be a vector of consumption goods. On the other hand, people are called *other-regarding*, if they care about others' payoffs, as well as their own.

Suppose there are two players 1 and 2. Let x_i be player i's material payoff. Each player i has preferences over (x_1, x_2), which I assume are represented by a utility function $u_i(x_1, x_2)$. Note that if player i is purely self-interested, his utility function depends only on x_i.

Various other-regarding preferences (alternatively called interdependent preferences or social preferences) can be explained by the following specification: For $i = 1, 2$ and $j \neq i$,

$$u_i(x_1, x_2) = x_i + g_i(x_j - x_i)x_j \tag{1}$$

The main features of this specification are that it is additively separable in the player's own material payoff and his concern for the other, and the latter part is multiplicatively separable in the function of his relative payoff and the other's material payoff.[8]

Two simplest examples of other-regarding preferences of this form are *altruism* and *spite*. Player i is purely altruistic if $g_i(\cdot)$ is constant and positive, while he is purely spiteful

[7] See Camerer (2003, Section 2.8), Fehr and Schmidt (2003), and Sobel (2001) for extensive surveys.

[8] See Neilson (2002) and Segal and Sobel (1999) for relevant axiomatization.

if $g_i(\cdot)$ is constant and negative. A little more elaborate specification is to assume $g_i(x_j - x_i) = \overline{g}_i$ if $x_i \geq x_j$ and $g_i(x_j - x_i) = \underline{g}_i$ if $x_i < x_j$ with $\overline{g}_i > \underline{g}_i > 0$: Although each player exhibits altruism, he puts more weight on the other's payoff if he is "ahead" ($x_i \geq x_j$) than if he is behind ($x_i < x_j$).[9]

Another category of examples places further restrictions on (1) as follows:

$$u_i(x_1, x_2) = x_i + g_i(x_j - x_i) \tag{2}$$

Function $g_i(\cdot)$ is defined as

$$g_i(z) = \begin{cases} -\alpha_i v(z) & \text{if } z \geq 0 \\ -\beta_i v(-z) & \text{if } z \leq 0 \end{cases} \tag{3}$$

where $v(z)$ is defined for $z \geq 0$ with $v(0) = 0$ and is strictly increasing, and α_i and β_i are constants with $\alpha_i > 0$. That is, when player i is behind ($x_j - x_i > 0$), he prefers to reduce the inequality in payoffs between two. If in addition $\beta_i > 0$, the same is true when player i is ahead ($x_j - x_i < 0$). The player with this type of other-regarding preferences is called *inequity averse*. For example, Fehr and Schmidt (1999) assume the following piecewise linear utility function.[10]

$$u_i(x_1, x_2) = \begin{cases} x_i - \alpha_i(x_j - x_i) & \text{if } x_i \leq x_j \\ x_i - \beta_i(x_i - x_j) & \text{if } x_i \geq x_j \end{cases} \tag{4}$$

On the other hand, if $\beta_i < 0$, player i prefers to increase the difference in payoffs when he is ahead. This preference can be called *competitive* or *status-seeking*.[11]

---

[9] An example of such a model is Charness and Rabin (2002).

[10] Bolton and Ockenfels (2000) propose an alternative specification of inequity aversion, which is not in form (2). They write player i's utility function as $u_i(x_i, s_i)$ with

$$s_i = \begin{cases} \frac{x_i}{x_1 + x_2} & \text{if } x_1 + x_2 \neq 0 \\ \frac{1}{2} & \text{if } x_1 + x_2 = 0 \end{cases}$$

where $\partial u_i / \partial s_i(x_i, 1/2) = 0$ and $\partial^2 u_i / \partial s_i^2(x_i, s_i) < 0$. $u_i(x_i, s_i)$ is thus maximized at $s_i = 1/2$: Each player will sacrifice to move his share closer to the average if he is either below or above it. Some experimental evidence that compares Bolton and Ockenfels (2000) with Fehr and Schmidt (1999) is discussed in Camerer (2003, Section 2.8.5).

[11] This terminology follows Neilson and Stowe (2003).

In the theories of other-regarding preferences presented above, only the final monetary distribution among players matters. In this sense, these models are often called the distributional approach. The second approach to modeling other-regarding players pays attention to intentions behind behavior. In particular, intention-based reciprocal behavior is regarded as one of the most important other-regarding behavior. Reciprocal behavior is a response to actions, or intentions behind actions of others, and can be either positive or negative. Positive reciprocity implies that if actions by others benefit a player, or if he perceives their actions as kind, then he returns the kindness to make the others better off. Negative reciprocity implies that if the others' actions are harmful or are perceived as unkind, then he retaliates to make the others worse off.

A seminal paper in the intention-based approach is Rabin (1993). He proposes that a player's preferences depend on strategies of the players, his beliefs about the other's strategy, and his beliefs about the other's beliefs about his strategy:

$$u_i(\sigma_i, \sigma_j, \sigma_i') = v_i(\sigma_i, \sigma_j) + g_i(\sigma_i', \sigma_j) v_j(\sigma_i, \sigma_j) \tag{5}$$

where σ_i is player i's strategy, σ_j is player i's belief about player j's strategy choice, and σ_i' is player j's belief about what player j believes about player i's strategy choice.[12] Since beliefs enter into preferences, Rabin (1993) utilizes psychological game theory (Geanakoplos et al., 1989) and solves for equilibria in strategies and beliefs. Rabin's model is only for normal-form games, while Dufwenberg and Kirchsteiger (2002) and Falk and Fischbacher (2000) extend his theory to extensive-form games.

These equilibrium models are difficult to use in most applications, because they are very involved and in general have multiple equilibria. Charness and Rabin (2002), Cox and Friedman (2002), and Levine (1998) attempt to develop more tractable models in the spirit of intention-based reciprocity. For example, Levine (1998) assumes that player i differs in the extent (type) to which he cares about the others, and if his type is $\alpha_i > 0$, his preferences are represented by

$$u_i(x_1, x_2) = x_i + \frac{\alpha_i + \lambda \alpha_j}{1 + \lambda} x_j \tag{6}$$

where $0 \leq \lambda \leq 1$. This is another example of specification (1). The idea is that player i cares more about player j's material payoff if player j cares more about player i (α_j is higher).

[12] See Segal and Sobel (1999) for a related axiomatic approach to generating preferences that can reflect intention-based reciprocity.

In this paper, I adopt the distributional approach, and in particular, the theory of inequity aversion a la Fehr and Schmidt (1999) for the following reasons. First of all, it is simple and tractable. Second, it is not an ad hoc specification tailored for a particular finding. It is a parsimonious specification of other-regarding preferences, supported by psychological intuition, that can explain various types of experimental results with a single function. And it is already a well studied model. For example, an axiomatic foundation for this model has been developed (Neilson, 2002). Third, it is argued that "[i]n many situations, reciprocal persons and inequity averse persons behave in similar ways (Fehr and Fischbacher, 2002, p.C3)" and they seem to be more important than pure altruism and spitefulness. The inequity aversion model can be used as a "shortcut" for studying the effects of important intention-based reciprocal behavior. However, we should be aware that reciprocity and inequity aversion are distinct motives, and often intention matters more, in particular in the domain of punishing behavior, as suggested by recent evidence.[13]

3 The Principal-Agent Model with Other-Regarding Preferences

3.1 Benchmark: The Self-Interest Case

I analyze the effects of other-regarding preferences in the following simple but standard principal-agent framework. The principal hires an agent for engaging in a project. Both are assumed to be risk neutral. The agent chooses an action from $A = \{a_0, a_1\}$. When action a_i is chosen, the agent incurs private cost d_i. I assume $d_1 > d_0 = 0$ and denote $d_1 = d$ for simplicity. The project succeeds with probability p_i and generates benefit b_s for the principal, while it fails with probability $1-p_i$ and the benefit is b_f. I assume $1 > p_1 > p_0 > 0$ and $b_s > b_f = 0$, and denote $b_s = b$ for simplicity.

The outcome of the project is verifiable, and the principal offers a contract (w_s, w_f) in which w_s is the payment to the agent when the project succeeds and w_f is paid when it fails.

[13]See, for example, Fehr and Schmidt (2003) for a survey. Note that according to them, "the evidence also suggests that inequity aversion plays an additional, nonnegligible role (p.238)." Charness and Haruvy (2002) estimate and compare various theories using experimental data on gift exchange games. They argue that distributional concern as well as intention play a significant role in players' decisions.

I assume that the contract must satisfy the *limited liability constraints for the agent*:

$$w_j \geq 0, \quad j = s, f \tag{LL1}$$

I denote the set of feasible contracts by $C = \{(w_s, w_f) \mid (w_s, w_f)$ satisfies (LL1)$\}$.

The timing of the game is as follows. First, the principal offers a contract. The agent decides whether to accept or reject the contract. If rejected, the game ends and the agent receives the reservation utility $\overline{u}$ which I assume is zero.[14] After accepting the contract, the agent chooses an action. The outcome of the project then realizes and the transfer is made according to the contract. I assume that the contract is binding and cannot be renegotiated.

In the standard setting in which all the parties are purely self-interested, the payoffs to the principal and the agent are given as $u_P = b_j - w_j$ and $u_A = w_j - d_i$, respectively, for action a_i and outcome $j = s, f$.

Throughout the paper I assume that b is sufficiently large so that the principal prefers to implement a_1 to a_0. The incentive compatibility constraint and the participation constraint are, respectively, given as follows:

$$\Delta_p \Delta_w \geq d \tag{IC1}$$

$$w_f + p_1 \Delta_w \geq d \tag{PC1}$$

where $\Delta_w = w_s - w_f$ and $\Delta_p = p_1 - p_0$. The principal chooses $(w_s, w_f) \in C$ that minimizes the expected payment $w_f + p_1 \Delta_w$ subject to (IC1) and (PC1).

Consider any contract in C that satisfies (PC1) with equality: $w_f + p_1 \Delta_w = d$. It is a first-best solution if the action were enforceable. However, since the action is not verifiable, the contract must satisfy (IC1) as well. To this end, consider a first-best contract $(\overline{w}_s, 0)$ where $p_1 \overline{w}_s = d$. Substituting this contract into the left-hand side of (IC1) yields

$$\Delta_p \overline{w}_s = \Delta_p \left(\frac{d}{p_1} \right) < d$$

and hence no first-best contract satisfies (IC1). (IC1) thus must bind. The optimal second-best contract (w_s^*, w_f^*) is unique and is given by $w_f^* = 0$ and

$$w_s^* = \frac{d}{\Delta_p}.$$

[14]The results of the paper would continue to hold if $\overline{u} > 0$, although some care must be taken since the participation constraint would sometimes bind.

To show the uniqueness, suppose $(w_s, w_f) \neq (w_s^*, 0)$ solves the principal's problem. Since (IC1) must bind at optimum, $\Delta_w = w_s - w_f = w_s^*$ and $w_f > 0$ holds. Then the principal's expected payment is $w_f + p_1\Delta_w > p_1 w_s^*$, which contradicts the optimality of (w_s, w_f).

The benchmark results shown above are summarized in the following proposition.

Proposition 1 *Suppose both principal and agent are self-interested. The optimal contract is unique and is given by* $(w_s^*, 0)$. *The first-best solution cannot be implemented.*

There exists a moral hazard problem, and in order to mitigate the problem, the principal must provide higher-powered incentives and leave rents $p_1 w_s^* - d = p_0 d/\Delta_p > 0$ to the agent.[15]

3.2 Other-Regarding Agent

I now suppose that the agent is other-regarding and cares about the principal's material payoff as well as his own. I assume that the agent has the following utility function: For $i = 0, 1$ and $j = s, f$,

$$u_A = w_j - d_i - \alpha v\left(\max\{b_j - 2w_j, 0\}\right) - \alpha\gamma v\left(\max\{2w_j - b_j, 0\}\right) \tag{7}$$

where $\alpha \geq 0$ and γ are constants. This utility function is just a restatement of (2), and I assume $v(0) = 0$, $v'(z) > 0$ for all $z > 0$, and $\lim_{z\to\infty} v(z) = \infty$.

The important assumption behind (7) is that the agent compares his income w_j with the principal's $b_j - w_j$: In other words, it is assumed that the agent does not take into account the disutility of his action d_i. I later analyze the alternative specification in which the agent compares his "net" material payoff $w_j - d_i$ with the principal's.

Note that when $2w_j - b_j > 0$ holds, the principal's monetary payoff is smaller than the agent's, and hence the agent is "ahead," while the agent is "behind" when $b_j - 2w_j > 0$ holds. (7) can thus be rewritten as follows.

$$u_A = \begin{cases} w_j - d_i - \alpha\gamma v(2w_j - b_j) & \text{if } w_j \geq b_j - w_j \\ w_j - d_i - \alpha v(b_j - 2w_j) & \text{if } w_j \leq b_j - w_j \end{cases} \tag{8}$$

[15] If the principal implements a_0, the optimal contract is $w_s = w_f = 0$ and the principal's expected payoff is $p_0 b$. Since I am assuming that the principal prefers to implement a_1 to a_0, the following condition is assumed implicitly: $\Delta_p b > p_1 d/\Delta_p$.

Parameter α captures the extent to which the agent cares about the principal's material payoff. If $\gamma < 0$, the agent prefers to increase the difference in payoffs when he is ahead, and hence he is competitive or status-seeking. On the other hand, if $\gamma > 0$, the agent is inequity averse and his utility is decreasing in the difference in payoffs between the principal and the agent, whether the agent is behind or ahead. Based on experimental evidence, Fehr and Schmidt (1999) assume $\gamma \leq 1$, that is, the agent suffers from inequality more when he is behind than when he is ahead. At this point I do not impose this restriction. Note that Fehr and Schmidt (1999) characterize the utility function by two parameters α and $\beta = \alpha\gamma$. I use γ instead of β since by changing α I can examine how the extent to which the agent is other-regarding affects incentives.

To simplify the exposition, I assume $w_f = 0$. This assumption can be justified if the principal's limited liability constraints ($w_s \leq b$ and $w_f \leq 0$) are imposed. Alternatively, I will show later in this section that the optimal contract in fact satisfies $w_f = 0$ if an additional assumption is made instead of the principal's limited liability. Since the agent does not suffer from inequality when the project fails, his other-regarding preferences are concerning how the benefit from the successful project is divided between the principal and the agent. The incentive compatibility constraint is then written as follows.

$$w_s - \alpha\gamma v(2w_s - b) \geq \frac{d}{\Delta_p} \qquad \text{if } w_s \geq \frac{1}{2}b \tag{IC2a}$$

$$w_s - \alpha v(b - 2w_s) \geq \frac{d}{\Delta_p} \qquad \text{if } w_s \leq \frac{1}{2}b \tag{IC2b}$$

where (IC2*a*) represents the incentive compatibility constraint when the agent is *a*head, while (IC2*b*) is the constraint when he is *b*ehind.

I first obtain a necessary and sufficient condition for the existence of a contract satisfying (IC2b).

Lemma 1 *There exists a contract that satisfies (IC2b) if and only if*

$$\frac{1}{2}\Delta_p b \geq d \tag{9}$$

holds.

Proof Since the left-hand side of (IC2b) is increasing in w_s, the existence is guaranteed if (IC2b) is satisfied at $w_s = b/2$, that is, if (9) holds. Conversely, if $w_s \leq b/2$ satisfies (IC2b), then

$$\frac{d}{\Delta_p} \leq w_s - \alpha v(b - 2w_s) \leq \frac{1}{2}b$$

holds. Q.E.D.

Define $\hat{w}_s$ implicitly by

$$\hat{w}_s - \alpha v(b - 2\hat{w}_s) = \frac{d}{\Delta_p}. \tag{10}$$

Note that $\hat{w}_s$ satisfies $\hat{w}_s \leq b/2$ if (9) holds. Since I focus on the implementation of action a_1, I maintain (9) for most of the analysis and state it as an assumption. The case in which (9) does not hold will be discussed later in this section.

Assumption 1 The profit to the principal from the successful project is so large that (9) holds.

The first proposition presented below states that $(\hat{w}_s, 0)$ is the optimal contract under Assumption 1.

Proposition 2 *Suppose Assumption 1 holds. (i) $(\hat{w}_s, 0)$ is optimal. (ii) The principal's expected payment under the optimal contract is increasing in α. If (9) holds with strict inequality, it is strictly increasing in α.*

Proof (i) It is sufficient to show that $(\hat{w}_s, 0)$ satisfies the participation constraint:

$$w_s - \alpha v(b - 2w_s) \geq \frac{d}{p_1} \tag{PCb}$$

By definition, $\hat{w}_s - \alpha v(b - 2\hat{w}_s) = d/\Delta_p > d/p_1$. (PCb) is hence satisfied. (ii) The principal's expected payment is $p_1\hat{w}_s$. It is thus sufficient to show that $\hat{w}_s$ is increasing in α. Define the left-hand side of (10) by $f(\alpha)$. Then

$$\frac{\partial f(\alpha)}{\partial \alpha} = -v(b - 2\hat{w}_s) < 0$$

for $\hat{w}_s < b/2$, which implies that $\hat{w}_s$ is strictly increasing in α. If (9) holds with equality, $\hat{w}_s$ does not depend on α. Q.E.D.

Proposition 2 implies that (i) similar to the pure self-interest case, at optimum the incentive compatibility constraint binds while the participation constraint does not, and hence the agent earns rents; and that (ii) the principal is worse off as the agent cares more about the principal's well-being. To understand why other-regarding preferences hurt the principal, remember that the left-hand side of (IC2b) is decreasing in α: Since the agent suffers from inequity when the project succeeds, the principal must pay more to satisfy the incentive compatibility constraint as the agent is more other-regarding.[16]

I next show that even if contracts with $w_f > 0$ are feasible, $(\hat{w}_s, 0)$ is uniquely optimal under an additional assumption.

Proposition 3 *Suppose that Assumption 1 holds and contracts with $w_f > 0$ are feasible. Then $(\hat{w}_s, 0)$ is the unique optimal contract if*

$$1 \geq 2\alpha\gamma v'(z) \tag{11}$$

is satisfied for all $z > 0$.

Proof Suppose that the conclusion does not hold and (w_s, w_f) with $w_f > 0$ is optimal. It implies that

$$w_f + p_1\Delta_w \leq p_1\hat{w}_s \tag{12}$$

holds. Since $\hat{w}_s \leq b/2$ and $w_f > 0$, $w_s < b/2$ must hold, and (w_s, w_f) satisfies the incentive compatibility constraint

$$\Delta_w - \alpha v(b - 2w_s) + \alpha\gamma v(2w_f) \geq \frac{d}{\Delta_p}. \tag{13}$$

Combining (13) and (10) yields

$$\begin{aligned}\Delta_w + \alpha\gamma v(2w_f) &\geq \hat{w}_s + \alpha v(b - 2w_s) - \alpha v(b - 2\hat{w}_s)\\ &\geq \Delta_w + \frac{w_f}{p_1} + \alpha v(b - 2w_s) - \alpha v(b - 2\hat{w}_s).\end{aligned}$$

The second inequality is due to (12). Rearranging yields

$$-\frac{w_f}{p_1} + \alpha\gamma v(2w_f) \geq \alpha v(b - 2w_s) - \alpha v(b - 2\hat{w}_s). \tag{14}$$

The right-hand side of (14) is positive since (12) leads to $w_s < \hat{w}_s$. Thus if the left-hand side is nonpositive, (w_s, w_f) does not satisfy the incentive compatibility constraint, which is a contradiction. (11) provides such a condition.[17] Q.E.D.

[16] The result that the principal prefers lower α will continue to hold if $\overline{u} > 0$ and the participation constraint binds. In this case, the principal has to increase payments for the obvious reason that the participation constraint becomes harder to satisfy as the agent is more other-regarding.

[17] Actually a weaker condition $1 \geq 2p_1\alpha\gamma v'(z)$, or $1 \geq p_1\alpha\gamma v'(0)$ if $\gamma v''(z) \leq 0$, is sufficient.

A contract with $w_f > 0$ creates inequity even when the project fails: The agent is ahead since he receives w_f while the principal pays w_f. If the agent is inequity averse ($\gamma > 0$), this change strengthens his incentive to choose a_1 as shown in the incentive compatibility constraint (13). However, for this change to benefit the principal, she must decrease w_s to lower Δ_w as well as to raise inequity $b - 2w_s$ when the agent is behind: These changes bring negative effects on incentives. Condition (11) is sufficient for the negative effects to dominate. If condition (11) is violated, then $1 < 2\alpha\gamma v'(z)$ for some $z > 0$, which implies that the agent may be willing to transfer some payment back to the principal in order to make him less ahead of the principal. Since this seems to be implausible, condition (11) is I believe a reasonable one. In particular, (11) holds if the agent is status-seeking ($\gamma < 0$).

When Assumption 1 Does Not Hold

The results obtained so far are based on Assumption 1. I now examine the case where Assumption 1 does not hold: Assume instead

$$\frac{1}{2}\Delta_p b < d \tag{15}$$

only in this subsection. By Lemma 1, there is no contract $(w_s, 0)$ that satisfies the incentive compatibility constraint in the range $w_s \leq b/2$. The principal thus has to choose $w_s > b/2$ to satisfy (IC2a). The left-hand side of (IC2a) is increasing in w_s if condition (11) holds, which I assume throughout this subsection.[18] Then define w_s^+ implicitly by

$$w_s^+ - \alpha\gamma v(2w_s^+ - b) = \frac{d}{\Delta_p}. \tag{16}$$

Writing the left-hand side of (16) as $g(\alpha)$ and differentiating it with respect to α yield

$$\frac{\partial g(\alpha)}{\partial \alpha} = -\gamma v(2w_s^+ - b)$$

for $w_s^+ > b/2$, and hence w_s^+ is strictly increasing (decreasing) in α if $\gamma > 0$ (respectively $\gamma < 0$). Therefore as in the previous result in Proposition 2 (ii), the principal does not benefit from a more "fair-minded" agent if he is inequity averse. However, if the agent is status-seeking, the principal is better off as the agent is more other-regarding. Since $w_s^+ > b/2$, the

[18] If (11) does not hold, no contract $(w_s, 0)$ can satisfy (IC2a) and hence can implement a_1. The principal must increase w_f from zero in order to encourage the agent to choose a_1, by making the agent suffer from the increased inequity facing the unsuccessful project.

agent is ahead if the project succeeds, and hence the status-seeking preferences come into play. The status-seeking agent enjoys being ahead, and thus the principal can implement a_1 with a lower cost as the agent is more competitive. This incentive effect arises only when condition (15) holds.

3.3 Alternative Specification

I have so far assumed that the agent compares his income w_j to the principal's $b_j - w_j$. However, since the agent knows he incurs the private cost of action d_i, he may be concerned about inequity differently, depending on whether his action is a_0 or a_1.

In this subsection, I assume alternatively that the agent compares his "net" material payoff $w_j - d_i$ to the principal's payoff $b_j - w_j$. Will the principal want the agent to be more inequity averse in this alternative specification?

First suppose $b - w_s > w_s$. Assuming $w_f = 0$ for simplicity, the incentive compatibility constraint and the participation constraint are, respectively, written as follows:

$$\Delta_p w_s - (p_1 \alpha v(b - 2w_s + d) - p_0 \alpha v(b - 2w_s)) - (1 - p_1)\alpha v(d) \geq d \tag{IC3b}$$

$$p_1 w_s - p_1 \alpha v(b - 2w_s + d) - (1 - p_1)\alpha v(d) \geq d \tag{PC3b}$$

Two modifications should be noted. First, the inequity aversion term under the successful project $\alpha v(b - 2w_s + d)$ contains d. However, if the agent chooses a_0 instead of a_1, the cost of action is zero, and hence the change in the inequity aversion term is

$$p_1 \alpha v(b - 2w_s + d) - p_0 \alpha v(b - 2w_s).$$

This change is larger under the current specification than the corresponding change in (IC2b) that is equal to $\Delta_p \alpha v(b - 2w_s)$, since by choosing a_0 the agent can save the cost of action and reduce the disutility from inequity when he is behind. In other words, by choosing a_0, he will not feel as envious as by choosing a_1 when he is behind. The incentive compatibility constraint is thus harder to satisfy than before.

The second change is the new inequity averse term $\alpha v(d)$ when the project fails. Since the agent incurs the cost of action d, he is again behind, and suffers from inequity. Choosing a_0 instead relieves him from envy he suffers from if the project fails. This change again makes the incentive compatibility constraint more tight.

Therefore the incentive compatibility constraint and the participation constraint become more stringent under this alternative specification than under the original specification, and hence the principal is again worse off the more other-regarding the agent is: The previous results are reinforced.

Next suppose $b - w_s < w_s - d$. The second change remains to be $\alpha v(d)$ while the first change is given by $p_1\alpha\gamma v(2w_s - b - d) - p_0\alpha\gamma v(2w_s - b)$. Now choosing a_1 *reduces* the difference by d when the agent is ahead. If the agent is status-seeking ($\gamma < 0$), both changes again make the constraint harder to satisfy. On the other hand, if the agent is inequity averse ($\gamma > 0$), the first change works so as to *benefit* the principal by making the incentive compatibility constraint easier to satisfy. However, the effect of the second change remains. For example, when $v(z) = z$, the second change dominates and the incentive compatibility constraint becomes harder to satisfy if and only if $1 - p_1 > p_1\gamma$.[19]

3.4 Other-Regarding Principal

The other-regarding behavior of the player at the role of the principal is more subtle to identify. The principal may behave fairly either because she is fair-minded or because she anticipates that otherwise the agent would respond to hurt the principal. Experimental evidence on proposers' behavior in ultimatum games along with the use of dictator games shows that both explanations are likely to be valid (Forsythe et al., 1994).

To see the effects of other-regarding preferences at the side of the principal, suppose the principal's utility function is given as follows: For $j = s, f$,

$$u_P = \begin{cases} b_j - w_j - \pi\delta y(b_j - 2w_j) & \text{if } b_j - w_j \geq w_j \\ b_j - w_j - \pi y(2w_j - b_j) & \text{if } b_j - w_j \leq w_j \end{cases} \tag{17}$$

where $\pi \geq 0$ and δ are constants, $y(0) = 0$, and $y'(z) > 0$ for $z > 0$.

Suppose Assumption 1 holds. The optimal contract when the principal is purely self-interested is $(\hat{w}_s, 0)$. Since $\hat{w}_s \leq b/2$, the principal is ahead when the project succeeds. This contract is still optimal when the principal's preferences are represented by (17), if the

[19] The remaining case is $w_s > b - w_s > w_s - d$. While the effect of the second change is the same, whether or not the first change makes the incentive compatibility constraint harder to satisfy is difficult to tell.

principal is better off as the payment is lower:

$$1 - 2\pi\delta y'(b - 2w_s) > 0$$

The condition is for example satisfied for the status-seeking principal ($\delta < 0$), or for the inequity averse principal ($\delta > 0$) but $\pi\delta$ sufficiently close to zero.

On the other hand, if

$$1 - 2\pi\delta y'(b - 2w_s) < 0$$

holds, then the principal prefers to *increase* the payment for the agent in order to reduce inequity when she is ahead (up to $b/2$). However, increasing the payment beyond $b/2$ makes her behind, and hence she prefers to *lower* the payment. Therefore $w_s = b/2$ holds at optimum: Since the principal is so much inequity averse, the optimal contract attains precisely the equal division between $b_j - w_j$ and w_j for $j = s, f$.

4 Multiple Agents

4.1 The Model

In this section, I extend the model of the previous section to a multi-agent setting, and assume that the principal does not belong to the agents' reference group while each agent cares about the payoff to the other agents. The production technology is the same as that in the previous section. The principal hires two agents denoted by $n = 1, 2$, each of which engages in a project separately. Each agent chooses an action from $A = \{a_0, a_1\}$. Action a_0 costs $d_0 = 0$ while a_1 costs $d_1 = d > 0$ privately to the agent. When action a_i is chosen, the project succeeds with probability p_i and generates profit $b_s = b > 0$ while it fails with probability $1 - p_i$ and the profit is $b_f = 0$. I assume that action a_1 is more productive and $1 > p_1 > p_0 > 0$. There is no correlation.

The timing of the game is as follows. First, the principal offers a contract to the agents. The agents decide simultaneously whether to accept or reject the contract. If rejected by at least one agent, the game ends and each agent receives the reservation utility zero. After both agents accept the contract, they choose actions simultaneously. The outcomes of the projects then realize and the transfers are made according to the contract.

Since the outcome of each project is verifiable, let $w^n = (w^n_{jk})_{j,k=s,f}$ be the payment scheme offered to agent n, in which w^n_{jk} represents the payment to agent n when his outcome is j and the other agent's outcome is k $(j, k = s, f)$. They must satisfy the limited liability constraints

$$w^n_{jk} \geq 0, \qquad j, k = s, f \tag{LL2}$$

Let $C_n = \{w^n \mid w^n \text{ satisfies (LL2)}\}$ be the set of feasible contracts for agent n, and $C = C_1 \times C_2$.

Agent n's utility function is given as follows: For $i = 0, 1$, $j, k = s, f$, $n, m = 1, 2$, and $m \neq n$,

$$u^n = w^n_{jk} - d_i - \alpha_n v_n \left(\max\left\{w^m_{kj} - w^n_{jk}, 0\right\}\right) - \alpha_n \gamma_n v_n \left(\max\left\{w^n_{jk} - w^m_{kj}, 0\right\}\right) \tag{18}$$

where $\alpha_n \geq 0$, $v_n(0) = 0$, and $v'_n(z) > 0$ for $z > 0$. I assume $|\gamma_n| \leq 1$: Being behind by a certain amount changes the agent's utility at least as much as being ahead by an equal amount. Fehr and Schmidt (1999) assume this for inequity averse agents based on experimental evidence: Each inequity averse agent dislikes inequality at least as much when he is behind as when he is ahead. I extend this assumption to status-seeking agents: Each status-seeking agent likes to be ahead no better than he likes to avoid being behind. Although I believe it is a reasonable assumption similar to loss aversion, it does not play an essential role in what follows.

Note that the agents compare the payments from the principal for each pair of project outcomes. The implicit assumption behind this formulation is that each agent either observes or estimates correctly what the other agent is paid, while the other agent's choice of action is not a concern. Since the project outcomes are verifiable and I will soon assume symmetric agents, the former part of the assumption seems reasonable. The latter part may be questionable, however. One could argue that the agents are likely to expect the other agent's action correctly and compare their actions as well. Since whether or not actions enter into comparison is an empirical question, I will first study this simpler specification, and later the other specification in which the agents compare their net payoffs.

To simplify the analysis, I adopt the following assumptions.

Assumption 2 (a) The agents are symmetric: $\alpha = \alpha_1 = \alpha_2$, $\gamma = \gamma_1 = \gamma_2$, and $v(\cdot) = v_1(\cdot) = v_2(\cdot)$. (b) The principal chooses a symmetric contract satisfying $w^1 = w^2$. (c) $v(z) = z$ for all

$z \geq 0$. (d) $\alpha\gamma \leq 1$.

If Assumption 2 (d) fails to hold, an inequity averse agent, when he is ahead, will want to give up his income in order to reduce the inequality between his and the other agent's payoff, which seems to be implausible.[20] Agent 1's utility function is then rewritten as follows (Agent 2's utility function can be similarly rewritten).

$$u^1 = \begin{cases} w^1_{jk} - d_i - \alpha\gamma(w^1_{jk} - w^2_{kj}) & \text{if } w^1_{jk} \geq w^2_{kj} \\ w^1_{jk} - d_i - \alpha(w^2_{kj} - w^1_{jk}) & \text{if } w^1_{jk} \leq w^2_{kj} \end{cases} \tag{19}$$

As before, I assume that b is large enough for the principal to want to implement (a_1, a_1). She chooses a symmetric contract $(w^1, w^2) = (w, w) \in C$ to minimize the expected payments.

In the appendix, I show that it is without loss of generality to restrict contracts to those satisfying $w_{fs} = w_{ff} = 0$, that is, those which pay the least possible amount to each agent when his project fails. Then the incentive compatibility constraint and the participation constraint are, respectively, written as follows:

$$p_1 w_{ss} + (1 - p_1)w_{sf} + (p_1 - (1 - p_1)\gamma)\alpha w_{sf} \geq \frac{d}{\Delta_p} \tag{IC4}$$

$$p_1 w_{ss} + (1 - p_1)w_{sf} - (1 - p_1)\alpha(1 + \gamma)w_{sf} \geq \frac{d}{p_1} \tag{PC4}$$

When the project of an agent fails and the other's project succeeds, the former agent is behind and suffers αw_{sf} from inequity aversion. On the other hand, if his project succeeds and the other's fails, he is ahead and his equity concern is represented by $\alpha\gamma w_{sf}$.

I call w a *team contract* if $w_{ss} > w_{sf}$, while w is called a *relative performance contract* if $w_{ss} < w_{sf}$. Finally, if $w_{ss} = w_{sf}$, the contract for each agent depends on the outcome of his project only, and hence it is called an *independent contract*. The principal chooses a feasible contract that minimizes the expected payments $2p_1(w_{sf} + p_1(w_{ss} - w_{sf}))$ subject to (IC4) and (PC4).

Benchmark: The Self-Interest Case

Before deriving the optimal contract, I solve the optimal contract for the benchmark case of purely self-interested agents ($\alpha = 0$). It is easy to show that only the incentive compatibility

[20] Although not necessary for the results of the paper, $\alpha\gamma \leq 1/2$ may be a more reasonable assumption: Otherwise, the agent who is ahead would want to renegotiate ex post to transfer his rewards to the other agent.

constraint binds, and any feasible contract that solves

$$p_1 w_{ss} + (1 - p_1) w_{sf} = \frac{d}{\Delta_p}$$

is optimal and the principal's expected payment is $2p_1 d/\Delta_p$. Note, in particular, that the independent contract $w_{ss} = w_{sf} = d/\Delta_p$ is an optimal contract.

4.2 Analysis

Now consider other-regarding agents ($\alpha > 0$). First, suppose that the incentive compatibility constraint (IC4) binds.

$$p_1 w_{ss} + (1 - p_1) w_{sf} = \frac{d}{\Delta_p} + ((1 - p_1)\gamma - p_1)\alpha w_{sf} \tag{20}$$

Note that the larger the left-hand side of (20) is, the higher the principal's expected payment is. Therefore if $(1 - p_1)\gamma > p_1$, the principal prefers to set $w_{ss} > w_{sf} = 0$, while if $(1 - p_1)\gamma < p_1$, the principal wants to set $w_{ss} = 0 < w_{sf}$.[21] Define $\hat{w}_{ss}$ and $\hat{w}_{sf}$ as follows.

$$\hat{w}_{ss} = \frac{d}{\Delta_p} \frac{1}{p_1} \tag{21}$$

$$\hat{w}_{sf} = \frac{d}{\Delta_p} \frac{1}{(1 - p_1) + \alpha(p_1 - (1 - p_1)\gamma)} \tag{22}$$

I analyze two cases separately.

Case 1: $(1 - p_1)\gamma > p_1$

The best contract among those under which the incentive compatibility constraint binds is $(w_{ss}, w_{sf}) = (\hat{w}_{ss}, 0)$. By (PC4) and $p_1 > \Delta_p$, this contract satisfies the participation constraint, and hence it is the optimal contract. The expected payment is $2p_1 d/\Delta_p$, which is equal to the expected payment under the optimal contract for the purely self-interested agents.

Case 2: $(1 - p_1)\gamma < p_1$

The incentive compatibility constraint binds at $(0, \hat{w}_{sf})$. I now derive the condition for $(0, \hat{w}_{sf})$ to satisfy the participation constraint. Substituting $(0, \hat{w}_{sf})$ into (PC4) yields

$$\frac{d}{\Delta_p} \frac{(1 - p_1)(1 - \alpha(1 + \gamma))}{(1 - p_1) + \alpha(p_1 - (1 - p_1)\gamma)} \geq \frac{d}{p_1}.$$

[21] If $(1 - p_1)\gamma = p_1$, any contract satisfying (20) is optimal.

A straightforward calculation yields the following necessary and sufficient condition for $(0, \hat{w}_{sf})$ to satisfy (PC4):

$$\frac{\ell_s}{\ell_f} \leq \frac{1 - \alpha\gamma}{\alpha} \tag{23}$$

where $\ell_s = p_1/p_0$ and $\ell_f = (1 - p_1)/(1 - p_0)$ are likelihood ratios. When conditions $(1 - p_1)\gamma < p_1$ and (23) hold, $(0, \hat{w}_{sf})$ is the optimal contract. I call this case Case 2a. The expected payment is calculated as

$$2p_1(1 - p_1)\hat{w}_{sf} = \frac{2d}{\Delta_p} \frac{p_1(1 - p_1)}{(1 - p_1) + \alpha(p_1 - (1 - p_1)\gamma)} \tag{24}$$

which is equal to $2p_1 d/\Delta_p$ at $\alpha = 0$, and is decreasing in α and increasing in γ.

When $(0, \hat{w}_{sf})$ fails to satisfy (23), the participation constraint (PC4) must bind:

$$p_1 w_{ss} + (1 - p_1)w_{sf} = \frac{d}{p_1} + (1 - p_1)\alpha(1 + \gamma)w_{sf} \tag{25}$$

Since the principal's expected payments is higher as the left-hand side is larger, the principal wants to choose w_{sf} as small as possible. However, $w_{sf} = 0$ does not satisfy the incentive compatibility constraint (IC4), and hence both (IC4) and (PC4) bind. Solving the simultaneous equations (25) and (20) provides the solution $(\overline{w}_{ss}, \overline{w}_{sf})$ as follows:

$$\overline{w}_{ss} = \frac{d}{\Delta_p} \frac{1 - p_1}{\ell_s p_1} \left(\frac{\ell_s}{\ell_f} - \frac{1 - \alpha\gamma}{\alpha} \right) \tag{26}$$

$$\overline{w}_{sf} = \frac{d}{\Delta_p} \frac{1}{\ell_s \alpha} \tag{27}$$

Note that $\overline{w}_{ss} > 0$ if (23) does not hold. The expected payment is calculated as

$$2p_1 \left(p_1 \overline{w}_{ss} + (1 - p_1)\overline{w}_{sf} \right) = \frac{2d}{\Delta_p} \frac{p_1(1 - p_1)}{\ell_s} \left(\frac{\ell_s}{\ell_f} + \gamma \right). \tag{28}$$

which is independent of α and increasing in γ. I call this case Case 2b. This case applies when $\alpha^{-1} - (\ell_s/\ell_f) < \gamma < p_1/(1 - p_1)$, which range exists if and only if $\alpha > (1 - p_1)p_0/p_1$.

The following proposition summarizes the results obtained.

Proposition 4 The optimal contract $w^* = (w^*_{ss}, w^*_{sf})$ is given as follows.

Case 1: $w^* = (\hat{w}_{ss}, 0)$ if $\gamma > p_1/(1 - p_1)$ holds. It is an extreme team contract. The expected payment does not depend on α or γ.

Case 2a: $w^* = (0, \hat{w}_{sf})$ if both $\gamma < p_1/(1 - p_1)$ and $\gamma \leq \alpha^{-1} - (\ell_s/\ell_f)$ hold. It is an extreme relative performance contract. The expected payment is decreasing in α and increasing in γ.

Case 2b: $w^* = (\overline{w}_{ss}, \overline{w}_{sf})$ if both $\gamma < p_1/(1 - p_1)$ and $\gamma > \alpha^{-1} - (\ell_s/\ell_f)$ hold. The expected payment is independent of α and increasing in γ.

Other-regarding preferences provide two incentive effects via the incentive compatibility constraint (IC4). An agent is behind if his project is unsuccessful while the other project succeeds. Since the agent faces an incentive to reduce the probability that he incurs disutility αw_{sf} from being behind, it brings a positive incentive effect. The second effect arises from the situation in which an agent is ahead. If his project succeeds while the other project fails, the inequity averse ($\gamma > 0$) agent suffers from being ahead, and hence he is discouraged to increase the probability of success. This second effect is negative for inequity averse agents. In Case 1 in Proposition 4, the second negative effect dominates, because (i) each agent is sufficiently averse in being ahead (γ large); or (ii) each project is relatively uncertain and uncontrollable ($p_1/(1 - p_1)$ small) so that the incentive effect from being ahead (due to the other's unsuccessful project) is more important than that from being behind (due to the other's successful project). Since the effects of inequity aversion should be minimized in this case, the extreme team contract is adopted. Note the important feature of extreme team contracts: They are "fair" in the sense that both agents are always paid exactly the same amount. However, exactly because of this feature, the principal's payoff turns out to be independent of the extent to which the agents care about each other's well-beings: She neither benefits nor suffers from the agents' other-regarding preferences.

If the project is relatively controllable, or the extent to which the agent is averse in being ahead is small, the first positive incentive effect dominates the second effect. This is Cases 2a or 2b in Proposition 4. The principal then wants to utilize this positive effect in her contract design by adopting a relative performance contract to generate the possibility of inequity. In particular, if the agents are status-seeking, the second incentive effect as well as the first effect are positive (they prefer being ahead). As long as the participation constraint does not bind (Case 2a), a tournament-like extreme relative performance contract emerges as an optimal contract even though there is no systematic shock and hence introducing "competition" does not benefit the principal if the agents are self-interested. Note that in Case 2a, as the agents are more other-regarding, the optimal "prize" of the tournament ($\hat{w}_{sf}$) becomes smaller and hence the the principal's expected utility increases.

However, the principal must compensate for the disutility from other-regarding preferences as shown in the left-hand side of (PC4). In Case 2b in which the the project outcome is very informative in terms of action choice, or the agents are sufficiently other-regarding, the participation constraint binds. The principal then ceases to offer the extreme relative performance contract and chooses a contract in which the agents are paid positive amounts whether the project succeeds or fails.

Now compare the optimal contract in Proposition 4 with the optimal contract for the purely self-interested agents ($\alpha = 0$). As I have already discussed above, there are usually many optimal contracts, and the independent contract $w_{ss} = w_{sf} = d/\Delta_p$ is optimal. However, a small amount of other-regarding preferences changes the optimal contract in an important way. The optimal contract is now generically unique, and independent contracts are no longer optimal (except for the insignificant cases), despite technological and stochastic independence. This result alerts the use of independent contracts in the analysis of agency models like ours, even if the agents were assumed to be purely self-interested.

4.3 Correlated Outcomes

One of the well known results from the principal-agent analysis with purely self-interested agents is the optimality of relative performance evaluation when the agents' performances are positively correlated (Holmstrom, 1982; Mookherjee, 1984). However, the main result in the previous section hints at the possibility that the optimal contract for the other-regarding agents may not be a relative performance contract in the correlated environment. To examine this possibility formally, in this section I extend the model to the case where the results of the projects are correlated.

The project of each agent either succeeds or fails contingent on not only his action and idiosyncratic shock (as in the original model) but also a common shock that affects both projects. The common shock is either good (with probability q) or bad (probability $1 - q$). If the common shock is good, then both projects succeed irrespective of the agents' actions. However, if the common shock is bad, then the project outcome of each agent depends on his action and idiosyncratic shock: His project succeeds with probability p_i and fails with probability $1 - p_i$ when his action is a_i, where p_i satisfy the same conditions as before. Each agent's project thus succeeds with probability $q + (1 - q)p_i$. Che and Yoo (2001) show in this setting the optimal contract is an extreme relative performance contract when the agents are purely self-interested.

Now consider other-regarding agents. Assuming $w_{fs} = w_{ff} = 0$, the incentive compatibility constraint and the participation constraint are, respectively, written as follows:

$$(1 - q)[p_1 w_{ss} + (1 - p_1)w_{sf} + (p_1 - (1 - p_1)\gamma)\alpha w_{sf}] \geq \frac{d}{\Delta_p} \tag{IC5}$$

$$(1 - q)[p_1 w_{ss} + (1 - p_1)w_{sf} - (1 - p_1)\alpha(1 + \gamma)w_{sf}] \geq \frac{d - qw_{ss}}{p_1} \tag{PC5}$$

The principal's expected payments are $2[qw_{ss} + (1 - q)p_1(p_1 w_{ss} + (1 - p_1)w_{sf})]$. Note that if $q = 0$, this model coincides with the previous one with independent outcomes.

Similarly to (21) and (22), define w^+_{ss} and w^+_{sf} as follows.

$$w^+_{ss} = \frac{d}{\Delta_p}\frac{1}{(1 - q)p_1}$$

$$w^+_{sf} = \frac{d}{\Delta_p}\frac{1}{(1 - q)[(1 - p_1) + \alpha(p_1 - (1 - p_1)\gamma)]}$$

Then $(w^+_{ss}, 0)$ is the extreme team contract and $(0, w^+_{sf})$ is the extreme relative performance contract such that the incentive compatibility constraint (IC5) binds. And it turns out that if (23) holds, $(0, w^+_{sf})$ satisfies (PC5) as well, and the principal's expected payments are the same as (24) that do not depend on q: By the extreme relative performance contract the principal can filter out the common shock. Note that since $\alpha = 0$ satisfies (23), $(0, w^+_{sf})$ with $\alpha = 0$ is the optimal contract for the purely self-interested agents as derived by Che and Yoo (2001).

It is also easy to find that the extreme team contract $(w^+_{ss}, 0)$ also satisfies (PC5) and the principal's expeted payments are given by

$$2(q + (1 - q)p_1^2)w^+_{ss} = \frac{2d}{\Delta_p}\frac{q + (1 - q)p_1^2}{(1 - q)p_1} \tag{29}$$

Note that the expected payments are increasing in q. And we know from the previous analysis that if $q = 0$ (no correlation) and $(1 - p_1)\gamma > p_1$ (Case 1), then the extreme team contract is optimal and hence the right-hand side of (29) is smaller than that of (24). Therefore, even in the positively correlated case, if q is sufficiently small, the extreme team contract is still optimal. This result is in contrast to the pure self-interest case.

4.4 Alternative Specification

I now study the alternative specification that the agents compare their material payoffs net of the costs of actions rather than their income. For example, employees who work closely may be able to monitor their actions each other, or even if actions are not mutually observable, they may be able to expect the other agent's action correctly, and hence their actions are likely to enter into comparison. Agent 1's utility function then changes as follows.

$$u^1 = \begin{cases} w^1_{jk} - d_i - \alpha\gamma(w^1_{jk} - d_i - w^2_{kj} + d_h) & \text{if } w^1_{jk} - d_i \geq w^2_{kj} - d_h \\ w^1_{jk} - d_i - \alpha(w^2_{kj} - d_h - w^1_{jk} + d_i) & \text{if } w^1_{jk} - d_i \leq w^2_{kj} - d_h \end{cases} \tag{30}$$

where $j, k = s, f$ and $h, i = 0, 1$, and i (h) is the index for agent 1's (2's) action.

To see how this alternative specification alters the results, suppose that both agents choose a_1. Their expected utility then does not change from the previous model, and hence the participation constraint is the same as (PC4). However, if one of the agents, say agent 1, chooses a_0 while the other chooses a_1, then the comparison is not between w^1_{jk} and w^2_{kj} but between w^1_{jk} and $w^2_{kj} - d$. The incentive compatibility constraint is thus summarized as follows (given symmetric schemes):

$$\begin{aligned} &p_1 w_{ss} + (1-p_1)w_{sf} + (p_1 - (1-p_1)\gamma)\alpha w_{sf} \\ &\qquad \geq (1-\alpha\gamma)\frac{d}{\Delta_p} + (1-p_0)p_1\alpha(1+\gamma)\frac{w_{sf}}{\Delta_p} \qquad \text{if } w_{sf} \leq d \end{aligned} \tag{31}$$

$$\begin{aligned} &p_1 w_{ss} + (1-p_1)w_{sf} + (p_1 - (1-p_1)\gamma)\alpha w_{sf} \\ &\qquad \geq (1-\alpha\gamma + (1-p_0)p_1\alpha(1+\gamma))\frac{d}{\Delta_p} \qquad \text{if } w_{sf} \geq d \end{aligned} \tag{32}$$

First consider an extreme team contract $(w_{ss}, 0)$. Since $w_{sf} = 0 < d$, (31) applies and the incentive compatibility constraint is simplified as follows.

$$p_1 w_{ss} \geq (1-\alpha\gamma)\frac{d}{\Delta_p} \tag{33}$$

Under the extreme team contract, the agent is always ahead by d if he deviates from a_1 and chooses a_0 while the other agent follows a_1. If he is status-seeking ($\gamma < 0$), he will enjoy this deviation, and hence the principal must provide stronger incentives to induce him to choose a_1 under the current specification than under the original specification.

On the other hand, if the agents are inequity averse ($\gamma > 0$), they are more strongly motivated to choose a_1 in order to avoid being ahead, and hence the principal can save the payments: The principal is better off because each agent feels guilty if he shirks while the other does not, and hence prefers avoiding such guilt. This nonpecuniary incentive substitutes the monetary one. And the principal's expected payments are now *decreasing* in α under the extreme team contract, and hence when the extreme team contract is adopted, the principal also prefers the agents to be more inequity averse. This result is in contrast to the result under the previous specification that the principal implementing extreme team contracts is indifferent in the agents' other-regarding preferences.

Furthermore, if the agents are sufficiently inequity averse, then the participation constraint becomes binding and hence the principal need not leave rents to them: The participation constraint under extreme team contracts is $p_1 w_{ss} \geq d/\Delta_p$, the right-hand side of which is at least as large as the right-hand side of the incentive compatibility constraint (33) if and only if $\alpha\gamma \geq 1/\ell_s$.

The new incentive benefit from the agents' caring about their actions also exists under relative performance contracts. However, the benefit is not as large under relative performance contracts as under team contracts. For example, suppose the extreme relative performance contract $(0, \hat{w}_{sf})$. A simple calculation shows $\hat{w}_{sf} > d$, and hence the incentive compatibility constraint (32) applies. The right-hand side of (32) is larger than that of (33) if $\gamma > -1$. Under $(0, \hat{w}_{sf})$, when the project of an agent fails and the other project succeeds, he is behind. When he is shirking, he compares his payoff zero to the other agent's payoff $\hat{w}_{sf} - d > 0$. It turns out that the extent to which he is behind is not as serious as in the same situation with unobservable actions (where the difference is $\hat{w}_{fs}$), and hence the new specification brings a negative incentive effect under the extreme relative performance contract. Note that under the team contract the agent choosing a_0 will never be behind and hence this negative effect is absent.

It is therefore likely that the agents' comparing their actions as well benefits team contracts more than relative performance contracts, and hence the extreme team contract is more likely to be optimal. In fact, I show the following results formally.

Proposition 5 (i) If $(1 - p_1)\gamma > p_1$ holds, the optimal contract is an extreme team contract. (ii) If $\alpha\gamma \geq 1/\ell_s$ holds, the extreme team contract such that the participation constraint binds is optimal.

Proof (i) Suppose that a contract (w_{ss}, w_{sf}) with $w_{sf} > 0$ is optimal. Define w'_{ss} by $p_1 w'_{ss} =$

$p_1 w_{ss} + (1 - p_1) w_{sf} - \epsilon$ and consider the extreme team contract $(w'_{ss}, 0)$, where $\epsilon > 0$. It is easy to see that the new contract satisfies the incentive compatibility constraint (31) or (32), and the participation constraint (PC4) for ϵ sufficiently small but positive. And the principal's expected payments are $2p_1^2 w'_{ss} < 2p_1(p_1 w_{ss} + (1-p_1)w_{sf})$, which contradicts the optimality of (w_{ss}, w_{sf}). (ii) The discussion preceding the proposition shows the condition given is necessary and sufficient for the extreme team contract under which the participation constraint binds to satisfy the incentive compatibility constraint (33) as well. The principal's expected payments are $2d$. Now consider an arbitrary contract (w_{ss}, w_{sf}). The principal's expected payments are

$$2p_1(p_1 w_{ss} + (1 - p_1)w_{sf}) \geq 2p_1 \left(\frac{d}{p_1} + (1 - p_1)\alpha(1 + \gamma)w_{sf} \right) \geq 2d$$

The first inequality follows from (PC4). The second inequality holds since $\gamma \geq -1$ (it is strict if $w_{sf} > 0$ and $\gamma > -1$). The extreme team contract is thus optimal. Q.E.D.

Proposition 5, along with Proposition 4, implies the following: (i) if an extreme team contract is optimal under the original specification, then the optimal contract is again an extreme team contract under the alternative specification; and (ii) there exists a range of parameter values in which although the optimal contract is not an extreme team contract under the original specification, the extreme team contract becomes optimal under the new specification. The conditions are given by $\alpha\gamma \geq 1/\ell_s$ and $(1 - p_1)\gamma < p_1$. It is possible for both of them to hold if $\alpha > p_0(1 - p_1)/p_1^2$. In this case, the nonpecuniary incentive is so strong that the incentive compatibility constraint can be ignored and the principal can do as well as if the agents were self-interested and their actions were enforceable.

5 Implications for Endogenous Preferences in Organizations

The introduction of other-regarding preferences into contract theory enables us to analyze how those preferences affect the optimal contract, and more importantly, what kinds of preferences the principal wants the agents to have. In other words, preferences can be part of the "contract" designed by the principal, as Perrow (1986) cited in the introduction section suggests.[22] In this regard, it is interesting to find that even Milgrom and Roberts (1992)

[22] Alternatively, the agent could change his preferences strategically. Rotemberg (1994) takes such an approach to studying organizational behavior.

take a sympathetic position: "Furthermore, important features of many organizations can be best understood in terms of deliberate attempts to change the preferences of individual participants to make these factors [such as altruism, exceedingly high regard for others' opinions of one's courage] more salient. As a result, organizationally desired behavior becomes more likely (Milgrom and Roberts, 1992, p.42)." Although the current paper is just a start and takes the agents' preferences as exogenous, the results provide some interesting implications for desirable preferences.

First, consider the single agent case. The main result is that the principal in general does not benefit from other-regarding preferences. If the benefit from the successful project is large, the principal prefers having a self-interested agent to an inequity averse or status-seeking agent.[23] In particular, the optimal incentive for the self-interested agent does not induce the other-regarding agent to choose a_1: more costly higher powered incentives are necessary.

If the benefit from the successful project is small, there may be a case in which the principal benefits from having a more status-seeking agent since he is motivated to choose a_1 to raise the possibility that he earns more than the principal: lower powered monetary incentives thus suffice.

Next consider the multi-agent case. Suppose that the principal can choose a contract as well as (α, γ) and the agents only compare what they are paid. The principal then wants the agents to have γ as small as possible and α sufficiently large. For example, if only inequity averse agents ($\gamma \geq 0$) are feasible, then the optimal preference for the principal is $\gamma = 0$ and $\alpha > \ell_f/\ell_s$, that is, the agent who feels sufficiently envious but does not feel guilty at all. If status-seeking agents are feasible as well, then the most competitive ($\gamma = -1$) and sufficiently other-regarding agent is the best. Note that Case 2b in Proposition 4 applies and

[23]Note that I have not covered purely altruistic agents. Actually modifying the range of the parameter values in the model allows the agent to be unconditionally altruistic (Rey Biel, 2003). I call the agent *altruistic* or *efficiency-seeking* if $\alpha < 0$ and $\gamma \leq -1$: The agent's utility is then increasing in the principal's payoff as well as in his income. If $\gamma = -1$, he is purely altruistic in the sense that his utility is identical whether he is behind or ahead. If $\gamma < -1$, he emphasizes his income more and the principal's payoff less when he is behind than when he is ahead. It is easy to show that the principal benefits from a more efficiency-seeking agent (with higher $|\alpha|$) for the nonpecuniary incentive to choose a_1 enables the principal to save the monetary incentive. And if the agent is sufficiently efficiency-seeking, no monetary incentive is needed and the participation constraint binds.

hence the optimal contract is neither extreme team nor extreme relative performance.

One reason why such a competitive preference pattern is desirable is a lack of productive interaction among agents in the model. However, the analysis reveals another interesting reason. If the principal can change the agents' preferences such that they compare their actions as well as the payments, she prefers implementing the extreme team contract for the agent with sufficiently inequity averse preferences (such that $\alpha\gamma \geq 1/\ell_s$ as in Proposition 5). Thus even though the agents work independently, implementation of a fair, team-based pay scheme may benefit the principal if she can change their preferences in deliberate manners.

6 Concluding Remarks

I have argued that incentives and other-regarding preferences interact in nontrivial ways. When an agent cares about the principal's well-being, she is in general worse off by having a more inequity averse agent. When there are multiple symmetric agents who care about each other's well-being, the principal can optimally exploit their other-regarding nature by designing an appropriate interdependent contract such as a "fair" team contract or a relative performance contract that creates inequality when their performance outcomes are different. The optimal contract depends on the nature of the agents' other-regarding preferences.

I believe behavioral contract theory be a fruitful approach to issues in organization. The approach taken in this paper can shed light on issues on endogenous preferences, as suggested by sociologists and even organizational economists. The future research should deal with organizational contexts or problems more explicitly to ask how various aspects of organizations affect the members' preferences. Another promising research theme is how members with various preferences should be grouped. To this purpose, extending the analysis to the cases not considered in the current paper, such as those of productive externalities, asymmetric agents, and more than two agents, is high on the list.

References

J. N. Baron (1988), “The Employment Relation as a Social Relation,” *Journal of the Japanese and International Economies.* 2:492–525.

J. N. Baron and D. M. Kreps (1999), *Strategic Human Resources: Framework for General Managers.* New York, NY: John Wiley&Sons.

T. F. Bewley (1999), *Why Wages Don’t Fall During a Recession.* Cambridge, MA: Harvard University Press.

G. E. Bolton and A. Ockenfels (2000), “ERC: A Theory of Equity, Reciprocity, and Competition,” *American Economic Review.* 90:166–193.

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

G. Charness and E. Haruvy (2002), “Altruism, Equity and Reciprocity in a Gift-Exchange Experiment: An Encompassing Approach,” *Games and Economic Behavior.* 40:203–231.

G. Charness and M. Rabin (2002), “Understanding Social Preferences with Simple Tests,” *Quarterly Journal of Economics.* 117:817–869.

Y.-K. Che and S.-W. Yoo (2001), “Optimal Incentives for Teams,” *American Economic Review.* 91:525–541.

J. C. Cox and D. Friedman (2002), “A Tractable Model of Reciprocity and Fairness,” `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=376981`.

M. Dufwenberg and G. Kirchsteiger (2002), “A Theory of Sequential Reciprocity,” mimeo.

F. Englmaier and A. Wambach (2002), “Contracts and Inequity Aversion,” `http://papers.ssrn.com/sol3/papers.cfm?abstract_id=358325`.

A. Falk and U. Fischbacher (2000), “A Theory of Reciprocity,” Working Paper No. 6, Institute for Empirical Research in Economics, University of Zurich.

E. Fehr and U. Fischbacher (2002), “Why Social Preferences Matter—The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives,” *Economic Journal.* 112: C1–C33.

E. Fehr, A. Klein, and K. M. Schmidt (2001), "Fairness, Incentives and Contractual Incompleteness," http://www.vwl.uni-muenchen.de/ls_schmidt/research/disc/abstracts/abst_fici.htm.

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

E. Fehr and K. M. Schmidt (2003), "Theories of Fairness and Reciprocity: Evidence and Economic Applications," In M. Dewatripont, L. P. Hansen, and S. J. Turnovsky (eds.), *Advances in Economics and Econometrics: Theory and Applications, Eight World Congress, Volume I.* Cambridge, UK: Cambridge University Press. ch. 6, pp. 208–257.

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

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

C. Grund and D. Sliwka (2002), "Envy and Compassion in Tournaments," http://papers.ssrn.com/sol3/papers.cfm?abstract_id=358363.

W. Güth, W. Klose, M. Königstein, and J. Schwalbach (1998), "An Experimental Study of a Dynamic Principal-Agent Relationship," *Managerial and Decision Economics.* 19: 327–341.

B. Holmstrom (1982), "Moral Hazard in Teams," *Bell Journal of Economics.* 13:324–340.

B. Holmstrom and P. Milgrom (1990), "Regulating Trade among Agents," *Journal of Institutional and Theoretical Economics.* 146:85–105.

B. Holmstrom and P. Milgrom (1991), "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization.* 7(Special Issue):24–52.

H. Itoh (1993), "Coalitions, Incentives, and Risk Sharing," *Journal of Economic Theory.* 60: 410–27.

C. Keser and M. Willinger (2000), “Principals’ Principles When Agents’ Actions Are Hidden,” *International Journal of Industrial Organization.* 18:163–185.

E. P. Lazear (1995), *Personnel Economics.* Cambridge, MA: The MIT Press.

D. K. Levine (1998), “Modeling Altruism and Spitefulness in Experiments,” *Review of Economic Dynamics.* 1:593–622.

P. Milgrom and J. Roberts (1992), *Economics, Organization and Management.* Englewood Cliffs, NJ: Prentice-Hall.

D. Mookherjee (1984), “Optimal Incentive Schemes with Many Agents,” *Review of Economic Studies.* 51:433–446.

H. R. Nalbantian and A. Schotter (1997), “Productivity Under Group Incentives: An Experimental Study,” *American Economic Review.* 87:314–341.

W. S. Neilson (2002), “An Axiomatic Characterization of the Fehr-Schmidt Model of Inequity Aversion,” mimeo.

W. S. Neilson and J. Stowe (2003), “Incentive Pay for Other-Regarding Workers,” mimeo.

C. Perrow (1986), *Complex Organizations: A Critical Essay, Third Edition.* New York, NY: Random House.

J. Pfeffer (1997), *New Directions for Organization Theory: Problems and Prospects.* New York, NY: Oxford University Press.

M. Rabin (1993), “Incorporating Fairness into Game Theory and Economics,” *American Economic Review.* 83:1281–1302.

M. Rabin (2002), “A Perspective on Psychology and Economics,” *European Economic Review.* 46:657–685.

P. Rey Biel (2003), “Inequity Aversion and Team Incentives,” mimeo.

J. J. Rotemberg (1994), “Human Relations in the Workplace,” *Journal of Political Economy.* 102:684–717.

U. Segal and J. Sobel (1999), "Tit for Tat: Foundations of Preferences for Reciprocity in Strategic Settings," Discussion Paper 99-10, Department of Economics, UC San Diego.

J. Sobel (2001), "Interdependent Preferences and Reciprocity," mimeo.

J. Tirole (1992), "Collusion and the Theory of Organizations," In J.-J. Laffont (ed.), *Advances in Economic Theory: Sixth World Congress, Volume II.* Cambridge: Cambridge University Press. ch. 3, pp. 151–206.

Appendix

In this appendix I show that in the model in Section 4, it is without loss of generality to restrict attention to contracts with $w^n_{fs} = w^n_{ff} = 0$. To this end, consider a symmetric contract $(w, w) \in C$ in which at least one equality of $w_{fs} \geq 0$ and $w_{ff} \geq 0$ is strict. Note that to simplify the notations, I drop the superscripts from the contracts. If $w_{sf} \geq w_{fs}$, The incentive compatibility constraint and the participation constraint are, respectively, written as follows:

$$(w_{sf} - w_{ff}) + p_1(w_{ss} + w_{ff} - w_{sf} - w_{fs}) + (p_1 - (1 - p_1)\gamma)\alpha(w_{sf} - w_{fs}) \geq \frac{d}{\Delta_p} \quad \text{(A1)}$$

$$w_{ff} + p_1(w_{sf} - w_{ff}) + p_1(w_{fs} - w_{ff}) + p_1^2(w_{ss} + w_{ff} - w_{sf} - w_{fs}) \\ -(1 - p_1)p_1\alpha(1 + \gamma)(w_{sf} - w_{fs}) \geq d \quad \text{(A2)}$$

Similarly, if $w_{sf} < w_{fs}$, they are written as follows:

$$(w_{sf} - w_{ff}) + p_1(w_{ss} + w_{ff} - w_{sf} - w_{fs}) + (p_1\gamma - (1 - p_1))\alpha(w_{fs} - w_{sf}) \geq \frac{d}{\Delta_p} \quad \text{(A3)}$$

$$w_{ff} + p_1(w_{sf} - w_{ff}) + p_1(w_{fs} - w_{ff}) + p_1^2(w_{ss} + w_{ff} - w_{sf} - w_{fs}) \\ -(1 - p_1)p_1\alpha(1 + \gamma)(w_{fs} - w_{sf}) \geq d \quad \text{(A4)}$$

The expected payment to each agent is

$$W = w_{ff} + p_1(w_{sf} - w_{ff}) + p_1(w_{fs} - w_{ff}) + p_1^2(w_{ss} + w_{ff} - w_{sf} - w_{fs}). \quad \text{(A5)}$$

Denote by $C^* \subset C$ the set of feasible contracts that satisfy the incentive compatibility constraint and the participation constraint.

Lemma A1 For a given contract $(w, w) \in C^*$ with $w_{sf} < w_{fs}$, there exists a contract $(w', w') \in C^*$ such that $w'_{sf} > w'_{fs}$ holds and the principal's expected payment under (w', w') is the same as that under (w, w).

Proof Define the new contract by $w'_{ss} = w_{ss}$, $w'_{ff} = w_{ff}$, $w'_{sf} = w_{fs}$, and $w'_{fs} = w_{sf}$. Obviously the new contract satisfies $w'_{sf} > w'_{fs}$ and the participation constraint. And the incentive compatibility constraint (A1) is satisfied because (w, w) satisfies (A3) and $p_1 - (1 - p_1)\gamma \geq p_1\gamma - (1 - p_1)$ holds for $\gamma \leq 1$. Finally, the expected payment under the new contract is equal to W. Q.E.D.

By Lemma A1, from now on I focus on contracts with $w_{sf} \geq w_{fs}$. Define a new pay scheme to each agent, $\hat{w} = (\hat{w}_{jk})$, that satisfies $\hat{w}_{fs} = \hat{w}_{ff} = 0$, $\hat{w}_{sf} = w_{sf} - w_{fs}$, and

$$p_1 \left(\hat{w}_{sf} + p_1(\hat{w}_{ss} - \hat{w}_{sf}) \right) = W. \tag{A6}$$

Condition (A6) implies that the principal's expected payment under the new contract $(\hat{w}, \hat{w})$ is equal to that under (w, w). It is easy to show $\hat{w}_{ss} > 0$.

I next show that the new contract $(\hat{w}, \hat{w})$ satisfies the participation constraint

$$p_1 \left(\hat{w}_{ss} + p_1(\hat{w}_{ss} - \hat{w}_{sf}) \right) - p_1(1 - p_1)\alpha(1 + \gamma)\hat{w}_{sf} \geq d \tag{A7}$$

The left-hand side is equal to

$$W - p_1(1 - p_1)\alpha(1 + \gamma)(w_{sf} - w_{fs})$$

which is the left-hand side of (A2). Thus by (A2), (A7) holds and the new contract satisfies the participation constraint.

Finally, the incentive compatibility constraint is written as follows.

$$\hat{w}_{sf} + p_1(\hat{w}_{ss} - \hat{w}_{sf}) + (p_1 - (1 - p_1)\gamma)\alpha\hat{w}_{sf} \geq \frac{d}{\Delta_p} \tag{A8}$$

By (A6) and $\hat{w}_{sf} = w_{sf} - w_{fs}$, the left-hand side of (A8) is equal to

$$\frac{W}{p_1} + (p_1 - (1 - p_1)\gamma)\alpha(w_{sf} - w_{fs}). \tag{A9}$$

Since $w_{fs} > 0$ or $w_{ff} > 0$ holds, it is easy to show

$$\frac{W}{p_1} > (w_{sf} - w_{ff}) + p_1(w_{ss} + w_{ff} - w_{sf} - w_{fs}). \tag{A10}$$

Therefore by (A10), (A9), and (A1), the new contract satisfies the incentive compatibility constraint (A8). The result is summarized in the following proposition.

Proposition A1 *For a contract $(w, w) \in C^*$ that satisfies $w_{fs} > 0$ or $w_{ff} > 0$, there exists a contract $(\hat{w}, \hat{w}) \in C^*$ such that $\hat{w}_{fs} = \hat{w}_{ff} = 0$ and the principal's expected payment is equal to that under (w, w).*