TI 2010-008/1
Tinbergen Institute Discussion Paper

Gift-Exchange, Incentives, and Heterogeneous Workers

Arjan Non

Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute.

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Gift-Exchange, Incentives, and Heterogeneous Workers*

Arjan Non†

May 31, 2011

Abstract

By incorporating reciprocity in an otherwise standard principal-agent model, I investigate the relation between monetary gift-exchange and incentive pay, while allowing for worker heterogeneity. I assume that some, but not all, workers care more for their principal when they are convinced that the principal cares for them. The principal can be egoistic or altruistic. Absent worker heterogeneity, an altruistic principal signals his altruism by offering relatively weak incentives and a relatively high expected total compensation. However, the latter is not always required to credibly signal altruism. Furthermore, since some workers do not reciprocate the principal's altruism, the principal may find it optimal to write a contract that simultaneously signals his altruism and screens reciprocal worker types. Such a contract is characterised by excessively strong incentives and a relatively high expected total compensation.

Keywords: reciprocity, gift-exchange, signaling game, incentive contracts, screening.

JEL-codes: D86, J41, M52, M55.

*I would like to thank Robert Dur, Sjoerd van den Hauwe, Vladimir Karamychev, Oke Onemu, Oleg Shchetinin, Joeri Sol, Paul Steffens, Jan Tichem, Joël van der Weele, three anonymous referees and seminar participants in Rotterdam and Maastricht for many useful comments. I am also grateful to participants of the workshop 'Social Relations and Incentives in the Workplace' held in Rotterdam, 2010, in particular to Samuel Lee. Finally, I thank participants of the CESifo Area Conference Employment and Social Protection, held in Munich 2011.

†Department of Economics, Erasmus University Rotterdam, and Tinbergen Institute. E-mail: janon@ese.eur.nl

1 Introduction

Economists generally recognise that human's pecuniary motives are not the only determinant of economic outcomes: other considerations play a significant role as well. One of these considerations that has received lots of attention recently is reciprocity, meaning that people are willing to promote the welfare of a kind person and reduce the welfare of an unkind person, even if it comes at a personal cost. The importance of reciprocal motivations in the workplace has been brought under the attention of economists by Akerlof's (1982) seminal paper on the gift-exchange hypothesis. He describes labor contracts as a gift-exchange between employers and employees, where employee's effort and employer's generous treatment of workers are reciprocal gifts.

Generous treatment of employees encompasses several aspects. Of all the aspects mentioned by Akerlof, the wage level has without doubt attracted most attention. The idea that workers reciprocate high wages by exerting more effort has been the subject of considerable empirical examination. Numerous laboratory experiments find a positive relation between employee's effort and the salary offered by the employer, although recent field studies are somewhat less supportive.[1]

In the light of these substantial efforts to empirically test the gift-exchange hypothesis, the limited amount of theoretical investigation is surprising. For instance, little is known about how employers optimally induce reciprocity when they do not only decide on the wage level, but also on the level of performance pay. Since monetary incentives are an essential element of labour contracts, the perceived generosity of a contract offer may well depend on its incentive intensity.

This paper studies this question by allowing for reciprocity in an otherwise standard principal-agent model. Thus, the principal decides on a base salary and a piece rate to compensate the worker. Workers are risk-averse and not protected by limited-liability. Although risk-aversion is a common assumption in agency models, its interaction with reciprocity is still an open question. Risk aversion is likely to affect the optimal composition of a gift, as workers are not neutral with respect to the variance of their expected income.

The modeling of reciprocity is inspired by Levine (1998)'s game-theoretic approach. This approach is based on the idea that to many workers, it matters whether their boss cares about them as a person or views them merely as a means to an end. The key assumption is that workers are conditionally

[1] An early experimental study is Fehr, Kirchsteiger and Riedl (1993). Fehr and Gächter (2000) survey the voluminous literature. See Dur et al. (2010) for an overview of recent field experiments.

altruistic: they care more about the principal's welfare when they sense that the principal cares for them. Specifically, I assume that the principal is either altruistic or selfish, and that conditionally altruistic workers are altruistic to the extent that they believe that the principal is an altruist. Thus, conditionally altruistic workers reciprocate a favor not because of the favor itself, but rather because it signals that the employer cares for them.

An advantage of this approach is that it offers a tractable way to distinguish between authentic and strategic kindness. When workers reciprocate a high wage by exerting more effort, an egoistic principal may want to pay a generous wage for strategic reasons: not because he cares about the workers' well-being, but because he wants to maximize his own profits. This concealed egoism puts the principal's generosity in a completely different light. Workers may therefore not automatically reciprocate high wages, but only when they sense that this kindness is genuine, rather than strategic. There is considerable experimental evidence that people do not only care about distributional outcomes, but are also concerned about the intentions behind an action (see e.g. Charness and Levine, 2007, and the references therein). Apart from the abundant experimental evidence, the relevance of intentions is also apparent in our criminal law system: sentence length depends not exclusively on the harm inflicted on the other party, but to a large extent on whether the harm was caused intentionally or by accident.

Although reciprocity is often considered to have a strong influence on labour contracts (see e.g. Bewley, 1999), a typical finding in laboratory experiments is that not all individuals are equally motivated by reciprocal tendencies. For example, in a three-person gift-exchange experiment by Gächter and Thöni (2010) about 25% of experimental subjects classify as 'egoistic', meaning that they are unwilling to exert effort regardless of the wage they and their colleague receive. Similar patterns are reported by Fischbacher et al. (2001) and Fischbacher and Gächter (2010) in the context of a public goods experiment. This heterogeneity in reciprocity raises the issue whether employers can possibly screen workers, and whether they find it profitable to do so. To investigate this issue, I assume throughout that not all workers are conditionally altruistic: some workers are exclusively motivated by their own material interest. I examine whether an altruistic principal can write a contract that signals his altruism, and at the same time screens conditionally altruistic workers.

The main focus of the paper is to characterize contracts that induce reciprocity, assuming the principal refrains from screening workers. The first result is that an altruistic principal offers weaker incentives than an egoistic principal, while at the same time increasing the base salary to ensure that he is viewed as an altruist. The reason is that a tight link between pay and

performance is not necessary when workers are convinced of the principal's kindness: strong incentives add little to the worker's productivity, but they do expose workers to (unnecessary) risk. Offering strong incentives is therefore suboptimal for a principal who really cares about the worker's well-being. Thus, weak incentives can be considered as part of the altruistic principal's gift to the worker.

The assumption of risk aversion is essential to this result. When workers would be risk-neutral, the principal can resolve all agency problems simply by equating the piece rate with the marginal product. By contrast, worker's risk aversion induces the principal to set the piece rate below the marginal product, and reciprocity can be helpful to further align the interest of the worker and the principal.

The second finding is related to the first and qualifies the standard result, namely that employers induce reciprocity by paying a relatively high expected total compensation. Because part of the altruistic principal's gift is a reduction in incentives, he need not necessarily pay a higher expected total compensation than an egoistic principal to signal his altruism. The reason is that, since an altruistic principal provides workers with little explicit incentives to exert effort, workers' effort may be relatively low, despite their altruistic feelings towards the principal. As a result, pretending to be an altruistic principal is not particularly profitable for an egoistic principal, implying that a relatively low total compensation suffices to distinguish both types. Thus, the interaction between incentives, risk-aversion, and considerations of strategic kindness may divert the usual positive relation between wages and effort. An altruistic principal pays a higher expected total compensation than an egoistic principal only if increased worker motivation leads to sizeable productivity gains, for instance because incentivizing workers via financial incentives is costly due to strong risk aversion.

The third result is that an altruistic principal may find it optimal to write a contract that signals his benevolent intentions, and simultaneously selects conditionally altruistic workers. Perhaps surprisingly, this is accomplished by setting excessively strong incentives and paying a relatively high expected total compensation. The reason for setting strong incentives is that conditionally altruistic workers exert more effort than egoistic workers, and hence have more to gain from an increase in pay-for-performance than egoistic workers. Thus, the paradox is that strong incentives are offered in order to attract the employees who need them the least.

These findings relate to a broad literature on the signaling value of incentives. In particular, the finding that in the absence of screening motives, an altruistic principal induces reciprocity by offering weaker incentives than an egoistic principal, is well in line with recent experimental evidence.

Several experimental studies find that incentives are hard to reconcile with reciprocity-induced voluntary cooperation. For instance, a laboratory experiment by Falk and Kosfeld (2006) shows that the implementation of a minimum effort requirement reduces effort, because a considerable group of individuals interpret such an action as a sign of distrust. The implementation of a fine in case effort does not meet a prescribed level has a similar effect, as shown by Fehr and Gächter (2002) and Fehr and List (2004). Their experimental evidence shows that, in the words of Fehr and List (2004, p. 743), "incentives based on explicit threats to penalize shirking backfire by inducing less trustworthy behavior". Ellingsen and Johannesson (2008) provide theoretical underpinnings for this behavior, arguing that control systems and incentives signal that the principal is not worth impressing. Likewise, Sliwka (2007) argues that incentives signal that selfish behavior is the social norm, which demotivates the conformistic agents in the population. These models crucially differ from the model presented here in that esteem and conformism drive the results, instead of reciprocity.

These findings extend to a more natural setting where the principal not only decides on controlling or trusting the agent, but also on a wage level. Falk and Kosfeld (2006) and Bartling et al. (2011) find that when employers have the opportunity to control the agent by limiting his effort discretion, employers either implement a control strategy, which consists of low effort discretion and low wages, or a trust strategy, which consists of high effort discretion and substantial wages. This suggests that paying a high wage and at the same time limiting a worker's effort discretion are conflicting signals. Controlling the worker therefore reduces the effectiveness of a wage gift.[2]

As noted above, this difficulty to reconcile incentives with voluntary cooperation is in line with my model. An important difference, however, is that I look at another type of incentives, namely piece rates instead of minimum effort requirements. As a result, in my model incentives are not necessarily viewed negatively. This allows altruistic employers to signal their altruism and at the same time screen workers by offering relatively strong incentives and a high expected total compensation.

[2]This phenomenon is also referred to as partial-crowding out (Fehr and Gächter, 2002): keeping the wage constant, voluntary cooperation is lower when (stronger) incentives are implemented, where voluntary cooperation is defined as the difference between actual and privately optimal effort. Fehr and Gächter (2002) and Bellemare and Shearer (2011) provide evidence in line with the partial-crowding hypothesis.

In my model, the reason for crowding is that a high piece rate diminishes the share of the marginal product that accrues to the principal, and therefore restricts the worker's opportunities to reciprocate the principal's favourable treatment. Hence, a tight link between effort and monetary reward reduces the principal's return to signaling altruism.

The paper proceeds as follows. The next section discusses the related literature. Section 3 sets out the model. Section 4 analyzes the observable types case, which serves as a benchmark for the analysis in section 5 where types are assumed to be unobservable. There, I first show that pooling equilibria do not exist, and then proceed to the analysis of separating equilibria. Finally, in section 6 I conclude and provide some avenues for further research.

2 Related literature

In economics, several authors have suggested ways of modeling reciprocity and underlying intentions, for example Rabin (1993), Levine (1998), Dufwenberg and Kirchsteiger (2004), Falk and Fischbacher (2006) and Battigalli and Dufwenberg (2009). Our model is based on Levine's approach because it provides a tractable and natural way to model the findings reported in organizational psychology and management: when employees infer that their manager or the organization cares about their well-being, they reciprocate with increased commitment, loyalty, and performance (see, for example, the reviews by Rhoades and Eisenberger 2002 and Cropanzano and Mitchell 2005). As pointed out by Gul and Pesendorfer (2010), this approach facilitates analysis and allows to distinguish between genuine kindness and instrumental kindness.

Despite the recent attention for modeling reciprocal behavior in the workplace, there are only few studies that investigate the relation between reciprocity and monetary incentives: Englmaier and Leider (2008), Arbak and Kranich (2005), and Bellemare and Shearer (2011). Compared to these studies, the key innovation is that I account for worker's risk-aversion as well as a concern for the intentions of the employer. The interaction between these assumptions plays a critical role when deriving the employer's optimal choice of base salary and bonus payments.

Englmaier and Leider (2008) incorporate reciprocity in a principal-agent model with risk averse agents. Their main finding is that incentive pay and reciprocal motivations are substitutes, which is qualitatively in line with my results. The key difference is that Englmaier and Leider largely ignore the importance of intentions, and instead assume that positive reciprocity is automatically induced when agents expect to receive a rent.

A working paper by Arbak and Kranich (2005) is closely related in the sense that incomplete information and Levine-type conditional altruism are key features of their modeling set-up. They assume that employers signal their type by increasing piece rates, but fail to analyze the case where employers can also resort to the base salary as a signaling tool.

In a related model, Bellemare and Shearer (2011) address this shortcoming. They find that employers induce reciprocity by increasing piece rates rather than by increasing the base salary. The main reason is that piece rates have a direct incentivizing effect, in addition to the effect of the reciprocity induced by the gift. This result follows from their assumption that workers are risk neutral and that a limited-liability constraint is always binding. Therefore, in the absence of a signaling motive piece rates are suboptimally low. As higher piece rates bring workers' incentives closer to the socially optimal level, the most efficient way to signal kindness is by increasing incentives. By contrast, in my paper stronger incentives expose workers to more risk, which is costly. Thus, assuming that workers are risk averse and unconstrained by limited-liability crucially affects the optimal composition of a wage gift.

Another paper that studies the relation between reciprocal motivations and incentive pay is Dur et al. (2010). However, their model does not allow for monetary gift-exchange, but focuses on social gift-exchange instead, meaning that the resources of the gift-exchange are non-monetary.[3] In particular, in Dur et al. (2010), the manager's gift exclusively consists of management attention. Another limitation is that they do not allow for strategic kindness of the principal. Both limitations are addressed by Dur (2009), but his model does not allow for incentive pay.

Another topic I address is how contracts can select reciprocal worker types and at the same time signal the principal's altruism. Interestingly, Kosfeld and Von Siemens (2009, 2011) investigate how employers use a combination of base salary and performance pay to screen reciprocal workers. An important difference is that they concentrate on reciprocity between colleagues who work in a team, rather than between worker and principal. The principal's contract offer therefore does not signal private information.

Sliwka (2007) does not study reciprocity, but he also studies contracts that signal the principal's private information and at the same time screen workers. As noted above, the main idea in his paper is that incentives signal that selfish behavior is the social norm, which demotivates the conformistic agents in the population. In addition to this signaling effect, incentives can also screen worker types when selfish and altruistic workers differ in their

[3] As argued by Akerlof (1982), manager's benevolent treatment of workers encompasses several aspects. Studies in management and organizational psychology distinguish between two broad categories: economic resources and socioemotional resources. Economic resources address "financial needs and tend to be tangible", whereas socioemotional resources address "social and esteem needs (and are often symbolic and particularistic)." (Cropanzano and Mitchell 2005, p. 881) In economics, monetary gift-exchange refers to the former, whereas social gift-exchange in the cited papers refers to the latter.

preferences over incentive intensity. Hence, the optimal decision whether to trust or to incentivize agents takes both the signaling and selection effect into account. An important difference with my analysis, however, is that in Sliwka the principal faces the binary decision whether to trust or incentivize workers, whereas in my model the principal has two continuous instruments (wage and base salary) at his disposal. Therefore, the principal has richer opportunities to simultaneously signal to and screen workers.[4]

3 The model

I consider a risk-neutral principal who is hiring one worker from a large population of workers. Workers are risk-averse and some of them are conditionally altruistic, meaning that the extent to which the worker is altruistic depends on the principal's altruism. The expected utility of a worker of type i is described by:

$$u_i = -\exp^{-r\left[b(e+\varepsilon)+s-\frac{1}{2}\theta e^2+\gamma_i(\alpha_j)E(\pi_j)\right]}. \tag{1}$$

This specification is widely used to describe the utility of risk-averse agents, where $r > 0$ captures the extent of risk-aversion. Production is simply given by effort e, but is prone to random shocks ε that are normally distributed with variance σ^2. Effort is non-contractible, but assuming that output can be observed, the worker earns a share b of observed output $(e+\varepsilon)$ and a base salary denoted by s. The worker's costs of effort are represented by $\frac{1}{2}\theta e^2$. Furthermore, $\gamma_i(\alpha_j)$ captures the strength of the worker's altruism towards his principal, which positively depends on the principal's altruism α_j. Note that the worker cares about the principal's expected payoff $E(\pi_j)$ instead of his actual payoff π_j. The reason is that it would be nonsensical to assume that the worker is risk-averse over the payoff of a risk-neutral principal. Workers differ in their altruism function $\gamma_i(\alpha_j)$, where $0 \leq \gamma_i(\alpha_j) < 1$. I distinguish between a worker's 'type' and a worker's 'altruism'. A worker's 'type' refers to his altruism function, whereas his 'altruism' refers to the outcome of the function $\gamma_i(\alpha_j)$. The typespace is specified below.

As is common in the literature, I remove the uncertainty on ε from the worker's utility function by deriving the certainty equivalent, which allows

[4] To the best of my knowledge, the only other paper that designs a contract that signals information to and at the same time screens the other contracting party, using two continuous instruments, is Soberman (2003).

for convenient transformation of the utility function into:[5]

$$E(u_i) = be + s - \frac{1}{2}\theta e^2 - \frac{1}{2}r\sigma^2 b^2 + \gamma_i(\alpha_j)E(\pi_j). \tag{2}$$

An increase in the bonus exposes the worker to more risk, which subtracts from the positive effect of the bonus on the worker's utility. To ensure that an increase in the bonus does not expose the worker to so much additional risk that his utility is lowered, I assume throughout that $r\sigma^2\theta \leq 1$.[6]

The expected payoff of a principal of type j is described by

$$E(\pi_j) = (1-b)e - s + \alpha_j E(u_i), \tag{3}$$

where $\alpha_j E(u_i)$ captures the altruistic feelings of the principal.[7] Analogous to the workers, I assume that $0 \leq \alpha_j < 1$. Principals differ in type α_j, implying that the worker's altruism depends on the specific match between the worker's and the principal's type.

Types are private information. There are two types of workers and principals, namely selfish (l) and (conditionally) altruistic (h) types. A selfish or egoistic principal has no altruistic feelings at all: $\alpha_l = 0$. An altruistic principal's care for the worker is denoted $\alpha_h \in (0,1)$. The prior probability that the principal is selfish is given by μ. Because the principal's type is unobservable, a worker's altruism depends on his beliefs concerning the principal's type. Workers update their beliefs after observing the principal's

[5] This standard transformation is only correct when there is no uncertainty on the principal's type, either because types are observable or because the worker puts all probability mass on a certain type. Uncertainty on the principal's type reduces the worker's welfare compared to the utility suggested by (2). Because, as we shall see, only separating equilibria exist, there is no uncertainty on the principal's type in equilibrium, and hence it is safe to ignore the worker's 'preference for certainty' in this domain for simplicity. Taking this preference into account would only strengthen the results, because deviation from the equilibrium strategy is less attractive when it leads to uncertainty about the principal's type.

[6] This assumption does not affect the main results. The mathematical condition is derived in the proof of lemma 4 given in the appendix.

[7] Obviously, the coefficients α_j and $\gamma_i(\alpha_j)$ depend on the units in which utility is measured, because utility must be measured in interpersonally comparable units. Therefore, the principal's payoff function only makes sense when the worker's utility is measured by the transformed utility function (2). Note that both players take the other's total payoff into account, so including the immaterialistic part of the payoff function. This assumption is not essential for the qualitative results and is not uncommon in the literature, see e.g. Becker (1974), Barro (1974), Bernheim and Stark (1988), Ley (1997), Sol (2010) and Dur and Sol (2010). Confining altruism to the material payoffs, as in Levine (1998), leads to a discrepancy between the worker's actual utility and what the principal believes the worker's utility to be. This is inconvenient analytically.

contract choice using Bayes' rule. When applicable, I rule out unreasonable beliefs by requiring that beliefs satisfy the intuitive criterion.[8]

Workers are either egoistic or conditionally altruistic/reciprocal. Specifically, I assume that an egoistic worker type never takes the principal's welfare into account regardless of the principal's altruism ($\gamma_l(\alpha_j) = 0$ for any α_j) and that a reciprocal worker type is completely egoistic when he believes that the principal is selfish ($\gamma_h(\alpha_l) = 0$). Thus, only a conditionally altruistic worker who believes that he is employed by an altruistic principal takes the welfare of the principal into account, i.e. $\gamma_h(\alpha_h) > 0$. I assume that a fraction δ of all workers in the population is selfish, whereas the remaining fraction $(1-\delta)$ is conditionally altruistic.

The timing of the game is as follows. First, the principal decides on a remuneration scheme (b, s). A worker accepts the contract if it yields him an expected utility of at least his reservation utility $\overline{u}$. Hence, as is common in principal-agent models, workers have no bargaining power. Also, I assume that $\overline{u}$ does not depend on $\gamma_i(\alpha_j)$: the value of a worker's outside option does not depend on his type. In case none of the workers accepts the contract, the principal's payoff is zero. Finally, the employed worker decides on his effort level e.

4 Analysis when types are observable

In this section, I assume that both players learn about each other's type before they make any decision. The reason for studying this case is that it yields some insights that will be valuable later on. I solve for a subgame perfect equilibrium using backward induction. The worker's effort choice follows from maximization of his utility function (2). The first-order condition is described by:

$$b - \theta e + \gamma_i(\alpha_j)(1 - b) = 0. \tag{4}$$

It is instructive to see what happens if $b = 0$ or if $b = 1$. If $b = 0$, the worker only exerts effort out of an altruism motive. By contrast, if the worker is residual claimant ($b = 1$), the worker's actions do not affect the principal's profits. Therefore, any worker type equates the marginal costs of effort with the marginal product ($\theta e = 1$), independent of his altruistic feelings. Rewriting the first-order condition (4) gives the worker's optimal

[8] Beliefs satisfy the intuitive criterion if, for all out-of-equilibrium actions, zero probability is assigned to player types that can only lose compared to their equilibrium payoff, see Cho and Kreps (1987).

effort choice e_i^*:

$$e_i^* = \frac{b + \gamma_i(\alpha_j)(1-b)}{\theta}. \tag{5}$$

It can easily be seen that effort increases in financial incentives for any b and in the worker's altruism as long as $b < 1$. Also, it is easily verified that altruism reduces the motivational effect of financial incentives ($\frac{de}{db}$) and vice versa: financial incentives reduce the responsiveness of effort towards altruism ($\frac{de}{d\gamma}$). The latter effect is intuitive: the larger the share of the marginal product that accrues to the worker, the smaller the share that accrues to the principal, hence the smaller the worker's possibilities to increase the principal's welfare. Therefore, the model predicts partial crowding-out of voluntary cooperation.[9] The negative effect of altruism on the motivational effect of financial incentives follows from the fact that the more the worker cares for his boss, the less he enjoys his bonus. In the extreme case that $\gamma_i(\alpha_j)$ approaches 1, the worker cannot be motivated by incentive pay because he cares about the principal's payoff as much as he cares about his own payoff.

The principal's choice of the optimal bonus b follows from maximization of his expected payoff, where he takes into account the worker's response to financial incentives and the worker's participation constraint:

$$\max_{s,b} \; E(\pi_j) = (1-b)e_i^* - s + \alpha_j E(u_i)$$
$$\text{s.t. } E(u_i) = be_i^* + s - \frac{1}{2}\theta e_i^{*2} - \frac{1}{2}r\sigma^2 b^2 + \gamma_i(\alpha_j)E(\pi_j) \geq \overline{u}.$$

Since the principal cares more about his own payoff than about the worker's utility ($0 \leq \alpha_j < 1$), it is not optimal to leave a rent to the worker. The principal thus reduces the base salary until the participation constraint binds. Inserting the base salary implied by the participation constraint into the objective function and differentiating with respect to b, we obtain the following first-order condition for optimal incentive provision:

$$\frac{de_i^*}{db}(1 - \theta e_i^*) - r\sigma^2 b = 0.$$

This condition elucidates the principal's trade-off. An increase in the bonus has one benefit and two costs. The benefit is that an increase in the bonus leads to additional effort, which contributes to the principal's payoff with the

[9] Voluntary cooperation is defined as the difference between actual effort (e_i^*) and privately optimal effort ($\frac{b}{\theta}$). Clearly, $\frac{\gamma_i(\alpha_j)(1-b)}{\theta}$ is decreasing in the bonus. Note that similar expressions can be found in Arbak and Kranich (2005) and Sliwka (2007).

size of the marginal product. However, because the worker's participation constraint is binding, the worker needs to be compensated for the additional cost of providing effort. Moreover, risk-averse workers need to be compensated for exposure to income uncertainty. We derive the payoff-maximizing bonus b^* by inserting (5) and its derivative into the first-order condition, which yields:

$$b^* = \frac{\left[1 - \gamma_i(\alpha_j)\right]^2}{\left[1 - \gamma_i(\alpha_j)\right]^2 + \theta r \sigma^2}. \tag{6}$$

Clearly, the bonus decreases in the worker's altruism $\gamma_i(\alpha_j)$. The reason is twofold. First, the more altruistic the worker, the smaller the motivational effect of financial incentives. In terms of the first-order condition, $\frac{de_i^*}{db}$ decreases in $\gamma_i(\alpha_j)$, hence reducing the marginal benefit of the bonus. The second reason is that the more altruistic the worker, the more effort he exerts, and consequently the higher his marginal costs of effort. Because the worker needs to be compensated for his costs of effort, it is more costly to stimulate effort further using financial incentives. Note that when the worker has no altruistic feelings towards his principal ($\gamma_i(\alpha_j) = 0$), we obtain the standard solution: the bonus is decreasing in the worker's risk aversion and in the variance of the error term.

A result that is harder to anticipate is the ambiguous total effect of an increase in $\gamma_i(\alpha_j)$ on effort. Given the level of financial incentives, worker's altruism has a positive effect on effort, but this is possibly more than offset by the corresponding decrease in financial incentives. Specifically, the total derivative $\frac{de_i^*}{d\gamma_i(\alpha_j)}$ is described by

$$\frac{de_i^*}{d\gamma_i(\alpha_j)} = \frac{r\sigma^2\left(\theta r\sigma^2 - \left[1 - \gamma_i(\alpha_j)\right]^2\right)}{\left(\left[1 - \gamma_i(\alpha_j)\right]^2 + \theta r\sigma^2\right)^2},$$

which is positive if $\theta r\sigma^2 > \left[1 - \gamma_i(\alpha_j)\right]^2$. Thus, an increase in $\gamma_i(\alpha_j)$ only has a positive effect on effort when workers are relatively risk-averse. Strong risk-aversion implies that the bonus is relatively small ($b^* < \frac{1}{2}$). As a result, additional effort has a relatively large effect on the principal's welfare, and hence the worker's effort is more responsive to changes in $\gamma_i(\alpha_j)$. By contrast, when workers are hardly risk-averse, the bonus is close to the marginal product, which restricts the worker's opportunities to benefit the principal. The negative effect of a reduction in incentives therefore outweighs the positive effect of an increase in worker's altruism.

The base salary is such that the worker's participation constraint is exactly satisfied:

$$s = \overline{u} - b^* e_i^* + \frac{1}{2}\theta e_i^{*2} + \frac{1}{2} r\sigma^2 b^{*2} - \gamma_i(\alpha_j) E(\pi_j). \tag{7}$$

As long as the principal's expected payoff is positive[10], total compensation is decreasing in the worker's altruism. The main reason is a compensating wage differential: the more utility workers derive from the non-monetary aspects of their job, the lower the required monetary compensation. Altruism leads to a rent that can be extracted when types are observable. A second reason is that the strength of financial incentives decreases in the worker's altruism, which is reflected in a lower risk-premium when workers are risk averse.

Worker's altruism has a positive effect on profits, for two reasons. First, because preferences are more aligned, altruistic workers are willing to exert more effort, given the strength of financial incentives (see (5)). Second, as more altruistic workers derive more utility from the non-monetary aspects of their job, they are willing to accept a lower base salary, keeping b constant (see (7)).

5 Analysis when types are unobservable

In the previous section we obtained some basic insights by assuming that types are observable. In this section I study the more interesting case where types are private information. Because neither the worker's type nor the principal's type is observable, contracts potentially have a dual role of both signaling the principal's type and screening worker's types. I consider four possible Perfect Bayesian equilibria, namely pooling and separating equilibria, where in each equilibrium the principal either screens or abstains from screening. In pooling equilibria, the equilibrium contract offer is independent of the principal's type. By contrast, in separating equilibria different types offer different contracts.

In the next two subsections, I will show that pooling equilibria do not exist. Then, I analyze a fully separating equilibrium where the principal signals his altruism, but abstains from screening worker types. Finally, I study a fully separating equilibrium where the altruistic principal writes a contract that signals his altruism and simultaneously screens worker types. I abstract from mixed strategy equilibria. Hence, throughout I refer to a fully separating equilibrium simply as separating equilibrium.

[10]Firms can only stay in business as long as expected profits are at least zero. It is easily verified that profits are increasing in the worker's altruism, so an altruistic principal obtains a positive expected payoff even when an egoistic principal earns zero profits in expectation.

5.1 Analysis: Pooling equilibrium

This section studies existence of a pooling equilibrium where the principal does not screen workers. I start the analysis by noting that any equilibrium contract, pooling or separating, has to satisfy four constraints. The first two constraints consist of each principal's incentive compatibility constraint (ICC): neither of the types should have an incentive to deviate. The other two constraints are the participation contraints (PC) of both worker types. In equilibria where principals do not screen workers, both worker types should be willing to accept the equilibrium contract(s). By contrast, screening of workers' types requires that the equilibrium contract is acceptable for conditionally altruistic workers only. In equilibrium, workers' beliefs should be correct. Thus, in a pooling equilibrium, workers believe that the principal is egoistic with probability μ.

In order to show that a pooling equilibrium does not exist, it is instructive to compare the profit functions of the egoistic and altruistic principal, see equation (3). The crucial difference between the two types is that the altruistic principal enjoys the expected utility of the worker, in addition to his monetary payoff. Whether a deviation from the equilibrium contract is profitable depends to a large extent on reciprocal workers' out-of-equilibrium beliefs. When reciprocal workers interpret deviation as a signal that the principal is egoistic, deviation is not profitable for both types of principals. By contrast, if deviation is interpreted as a signal of altruism, both types are willing to deviate.

The intuitive criterion, however, restricts the set of possible out-of-equilibrium beliefs. In particular, it rules out that the worker believes that a deviation comes from an egoistic principal when that deviation cannot be profitable for an egoist, while an altruist gains by such a deviation. Such a contract always exists. For instance, the altruistic principal can deviate by increasing the base salary with an amount equal to the additional revenues he gains when the worker believes with probability one that the contract is offered by an altruist. The egoist can only lose by offering such a contract. The altruistic principal, however, gains by offering such a contract, because his monetary payoff is unaffected, while at the same time the worker obtains a strictly higher utility, which is valuable for an altruist. Thus, the intuitive criterion dictates that such a deviation must come from an altruistic principal, implying that the altruistic principal always has an incentive to deviate. The following proposition summarizes the discussion above:

Proposition 1 *A pooling equilibrium where the principal does not screen workers, does not exist.*

5.2 Analysis: Pooling and screening equilibrium

Instead of pooling on a contract that is acceptable for both types of workers, principals may also pool on a contract that is acceptable for conditionally altruistic workers only. As noted above, screening requires that an egoistic worker's participation constraint is violated, while at the same time satisfying the other constraints. The worker's utility function (2) reveals that a reciprocal worker derives more utility from a given contract than a selfish worker, as long as he believes with positive probability that the principal is an altruist. Thus, the principal can screen workers by extracting (part of) reciprocal workers' rent. The possibility to screen workers, however, does not change the altruistic principal's incentive to break the proposed equilibrium:

Proposition 2 *A pooling equilibrium where the principal screens workers does not exist.*

Proof. See appendix. ■

The intuition is the same as in a pooling equilibrium where the principal does not screen workers: the altruistic principal always finds it profitable to reveal his type. He deviates by offering a contract that cannot be profitably offered by an egoistic principal, and by the intuitive criterion, such a deviation must be interpreted as a credible signal of his altruism. The altruistic principal enjoys the resulting increase in worker's utility, while earning the same (or slightly lower) monetary profits. A formal proof is provided in the appendix, as the proof is complicated by the screening constraint outlined above.

5.3 Analysis: Separating equilibria

This section studies separating equilibria: equilibria where the altruistic principal offers a contract that signals his altruism. As pointed out above, in a separating equilibrium the principal either screens or refrains from screening workers. I will distinguish between the two cases after laying out the common structure of these equilibria.

I start the analysis by deriving the egoistic principal's contract choice. Using the assumption that a selfish principal does not care about worker's utility, his expected payoff (see equation (3)) can be written as:

$$E(\pi_l) = (1 - b_l)\left[\delta e_{ll} + (1 - \delta)\, e_{hl}\right] - s_l, \tag{8}$$

where the subscript (l) is used to indicate that the remuneration scheme (b_l, s_l) is offered by a selfish principal. Similarly, e_{ij} denotes the effort of

a worker of type i who faces the incentive scheme offered by a principal of type j and consequently believes that he is employed by a principal of type j.[11] His effort choice is described by equation (5). Because beliefs should be correct in equilibrium, a worker of type i observing the contract (b_l, s_l) correctly believes that he is employed by a selfish principal, implying that his expected utility (2) is described by:

$$E(u_i) = b_l e_{il} + s_l - \frac{1}{2}\theta e_{il}^2 - \frac{1}{2} r\sigma^2 b_l^2 + \gamma_i(\alpha_l) E(\pi_l) \geq \overline{u}. \tag{9}$$

Clearly, because $\gamma_i(\alpha_l) = 0$ irrespective of a worker's type, both worker types exert the same effort and derive the same utility from accepting the selfish principal's equilibrium contract. This implies that both types require the same compensation to satisfy their participation constraint. Because in any separating equilibrium a selfish principal has no reason to screen or to signal his type, he does not distort his optimal contract choice compared to the case when types are observable. Thus, the following lemma applies:

Lemma 1 *In any separating equilibrium, the egoistic principal offers a bonus* $b_l = \frac{1}{1+r\sigma^2\theta}$ *and a base salary that exactly satisfies the worker's participation constraint:*

$$s_l = \overline{u} - b_l e_{il} + \frac{1}{2}\theta e_{il}^2 + \frac{1}{2} r\sigma^2 b_l^2. \tag{10}$$

Proof. To show that in any separating equilibrium the egoistic principal always finds it optimal to offer the contract specified above, first note that in any separating equilibrium screening makes no sense for the egoistic principal and is not even possible, as $\gamma_i(\alpha_l) = 0$ (see (9)). Thus, all possible equilibrium contracts satisfy the worker's PC as defined by (9). To show that the contract specified above is the unique equilibrium contract, consider any arbitrary separating equilibrium. As worker's beliefs are correct in equilibrium ($\gamma_i(\alpha_l) = 0$), the selfish principal cannot be confronted with a reduction in worker's altruism when deviating. Therefore, of all possible equilibrium contracts, the egoistic principal will always choose the contract that maximizes his expected payoff. The properties of this contract follow from the analysis when types are observable and $\gamma_i(\alpha_l) = 0$. ■

Although the egoistic principal's equilibrium contract choice is highly intuitive, it cannot be sustained for all out-of-equilibrium beliefs. For instance, a reduction in the base salary should be viewed as coming from an egoistic principal, otherwise the egoistic principal can gain by slightly reducing the

[11] Because in equilibrium beliefs are always based on the observed contract offer, this shorthand notation suffices to describe the equilibrium contracts.

base salary, hence selecting conditionally altruistic workers. It seems natural to assume that a reduction in the base salary does not lead to more positive beliefs. Similarly, a salary raise should not be regarded so optimistic that the raise is actually profitable. Formally:

Lemma 2 *In any separating equilibrium, the worker's out-of-equilibrium belief upon observing an alternative contract offer* (b', s')*, is defined as follows. Let* $E(\pi_l^*) = E(\pi_l(b_l, s_l))$ *denote the egoistic principal's expected payoff in any separating equilibrium. Let* $pr(\alpha_j = \alpha_h)$ *denote the worker's belief. There are two cases:*
1) When (b', s') *is such that* $E(\pi_l(b', s') \mid pr(\alpha_j = \alpha_h) = 1) \leq E(\pi_l^*)$*, the worker believes that* $pr(\alpha_j = \alpha_h) = 1$.
2) When (b', s') *is such that* $E(\pi_l(b', s') \mid pr(\alpha_j = \alpha_h) = 1) > E(\pi_l^*)$*, then* $pr(\alpha_j = \alpha_h)$ *should be such that* $E(\pi_l(b', s') \mid pr(\alpha_j = \alpha_h)) < E(\pi_l^*)$*. Such a belief always exists.*

Proof. The first condition directly follows from the requirement that beliefs satisfy the Intuitive Criterion. The second condition ensures that the egoistic principal has no profitable deviation. To show that such a belief always exists, notice that $pr(\alpha_j = \alpha_h) = 0$ implies that $E(\pi_l(b', s') \mid pr(\alpha_j = \alpha_h) = 0) < E(\pi_l^*)$. ■

Deriving the altruistic principal's equilibrium contract is more difficult. In a separating equilibrium, any contract that cannot be profitably offered by an egoistic principal is a possible equilibrium contract. However, it is easily verified that the intuitive criterion rules out all equilibria in which the altruistic principal offers a contract that does not maximize his expected payoff.[12] Therefore, of all possible equilibrium contracts, the altruistic principal always chooses the contract that maximizes his expected payoff.

To derive this contract, it is instructive to inspect the altruistic principal's payoff function. Assuming that the equilibrium contract (b_h, s_h) succeeds in credibly signaling the principal's altruism, rewriting equation (3) yields:

$$E(\pi_h) = (1 - b_h)[\delta e_{lh} + (1 - \delta) e_{hh}] - s_h + \alpha_h [\delta E(u_l) + (1 - \delta) E(u_h)]. \quad (11)$$

This equation shows that the altruistic principal's payoff positively depends on a worker's altruism for two reasons. First, as long as $b_h < 1$, a reciprocal worker will put more effort into his job than a selfish worker ($e_{hh} > e_{lh}$). Second, a worker's expected utility $E(u_i)$ increases in his altruism, which is

[12] The reason is that, since the equilibrium contract by definition satisfies the requirement that it cannot profitably be offered by an egoistic principal, the intuitive criterion dictates that the contract must be offered by an altruistic principal.

valuable for a principal who has altruistic feelings. For these two reasons, the altruistic principal may benefit from writing a contract that convinces reciprocal worker types that he is an altruist. In addition, he may find it profitable to screen workers, which means that $\delta = 0$. In order to credibly signal altruism, the equilibrium contract (b_h, s_h) should satisfy two incentive compatibility constraints: the selfish principal should have no incentive to mimic the altruist and vice versa:

$$(1 - b_l)e_{\delta l} - s_l \geq (1 - b_h)e_{\delta h} - s_h, \tag{ICC1}$$

$$(1-b_l)e_{\delta l}-s_l+\alpha_h\overline{u} \leq (1-b_h)e_{\delta h}-s_h+\alpha_h\left[\delta E\left(u_l\right)+(1-\delta)E\left(u_h\right)\right], \tag{ICC2}$$

where $e_{\delta l} = \delta e_{ll} + (1-\delta)\, e_{hl}$ and $e_{\delta h} = \delta e_{lh} + (1-\delta)\, e_{hh}$. The incentive compatibility constraints in case the altruistic principal screens workers assume that $\delta = 0$ on the right hand side of the inequalities. It is essential to note that when ICC1 is satisfied, ICC2 can only be satisfied if the difference $\alpha_h\left[\delta E\left(u_l\right)+(1-\delta)E\left(u_h\right)\right]-\alpha_h\overline{u}$ is large enough. This observation reveals why in equilibrium a principal with altruistic feelings is willing to engage in costly signaling: not because he earns a higher monetary payoff, but because he values workers' utility. The altruistic principal's monetary payoff is constrained by the egoistic principal's monetary payoff: an egoistic principal will imitate any contract that yields a higher monetary payoff than his own equilibrium contract. Therefore, any positive difference in effort $(e_{\delta h} - e_{\delta l})$ must be reflected in higher payments to the worker. Reasoning further along these lines, the following lemma does not come as a surprise:

Lemma 3 *The incentive compatibility constraint ICC1 is binding in any separating equilibrium, implying that ICC2 is slack.*

Proof. Suppose, per absurdum, that ICC1 is slack. The altruistic principal chooses the contract that maximizes his payoff. Depending on whether he screens or not, the contract satisfies the participation constraint of both workers or that of conditionally altruistic workers only. Since this contract maximizes his payoff, ICC2 is satisfied, but ICC1 is violated. To see this, recall that profits are increasing in the worker's altruism, because more altruistic workers exert more effort (see (5)) and are willing to accept a lower base salary (see (7)). Therefore, the egoistic principal finds it profitable to mimick the altruistic principal. Hence, ICC1 binds in any separating equilibrium. Because any equilibrium contract satisfies $E(u_i) \geq \overline{u}$ for the worker who accepts the contract, ICC2 is slack when ICC1 holds with equality. ■

Since ICC1 is binding, in any separating equilibrium the optimal contract will be a (b_h, s_h) −combination such that ICC1 holds with equality and that

ICC2 is slack.[13] I will now proceed to discuss the separating equilibrium where the principal does not screen.

5.3.1 Separating equilibrium: no screening

When the altruistic principal does not screen workers, his equilibrium contract should satisfy the participation constraint (PC) of both worker types. For ease of reference, the selfish worker's PC is described by

$$b_h e_{lh} + s_h - \frac{1}{2}\theta e_{lh}^2 - \frac{1}{2} r\sigma^2 b_h^2 \geq \overline{u}. \tag{PCL}$$

Assuming that the contract (b_h, s_h) credibly signals the principal's altruism, a reciprocal worker's PC is described by

$$b_h e_{hh} + s_h - \frac{1}{2}\theta e_{hh}^2 - \frac{1}{2} r\sigma^2 b_h^2 + \gamma_h(\alpha_h) E(\pi_h) \geq \overline{u}. \tag{PCH}$$

Comparison of these two constraints reveals that a reciprocal worker derives more utility from a given equilibrium contract (b_h, s_h) than a selfish worker. This implies that when a selfish worker's PC is satisfied, a reciprocal worker's PC is also satisfied. Because the equilibrium contract (b_h, s_h) by assumption satisfies the selfish worker's PC, the reciprocal worker's PC cannot be binding in the proposed equilibrium.

The problem is to find the contract that maximizes the altruistic principal's expected payoff, provided the four constraints outlined above are satisfied. Because PCH and ICC2 are both slack, the problem can be reformulated in a convenient way. As the altruistic principal's monetary payoff is constrained by the profits earned by the egoistic principal, he maximizes his total payoff by choosing a contract that exactly satisfies ICC1 and maximizes the expected utility of the worker. Denoting the egoistic principal's equilibrium profits as $E(\pi_l^*)$, the maximization problem can thus be formulated as:

$$\begin{aligned} \max_{s_h, b_h} \quad & E(\pi_h) = E(\pi_l^*) + \delta E\left(u_l\right) + (1-\delta) E(u_h) \\ \text{s.t.} \quad & s_h = (1-b_h) e_{\delta h} - E(\pi_l^*), && \text{(ICC1)} \\ \text{s.t.} \quad & b_h e_{lh} + s_h - \frac{1}{2}\theta e_{lh}^2 - \frac{1}{2} r\sigma^2 b_h^2 \geq \overline{u}. && \text{(PCL)} \end{aligned}$$

[13] Obviously, the altruistic principal should offer a different contract than the egoist, i.e. the contract $(b_h = b_l, s_h = s_l)$ is not feasible. As we shall see, this condition is always satisfied if $r\sigma^2 > 0$.

Intuitively, a reasonable conjecture is that maximization of worker's utility ensures that the selfish worker's PC is satisfied. For ease of exposition, I assume that PCL is satisfied and show afterwards that this conjecture is correct. This allows me to rewrite the problem by substituting $E(u_l)$, $E(u_h)$, and s_h, yielding the following:

$$\max_{b_h} \quad E(\pi_h) = \delta\left(e_{lh} - \frac{1}{2}\theta e_{lh}^2 - \frac{1}{2}r\sigma^2 b_h^2\right) + (1-\delta)\left(e_{hh} - \frac{1}{2}\theta e_{hh}^2 - \frac{1}{2}r\sigma^2 b_h^2 + \gamma_h(\alpha_h)E(\pi_h)\right).$$

Taking the derivative to b_h gives an insightful first-order condition:

$$\delta\left(\frac{de_{lh}}{db_h}(1-\theta e_{lh}) - r\sigma^2 b_h\right) + (1-\delta)\left(\frac{de_{hh}}{db_h}(1-\theta e_{hh}) - r\sigma^2 b_h\right) = 0.$$

The first-order condition is the same as in the observable types case, but weighted according to the prevalence of the two worker types. An altruistic principal thus offers the bonus that maximizes the expected surplus. Therefore, depending on the fraction of selfish workers in the population (δ), the payoff-maximizing bonus b_h lies between $b^* = \frac{[1-\gamma_h(\alpha_h)]^2}{[1-\gamma_h(\alpha_h)]^2+\theta r\sigma^2}$ and $b_l = \frac{1}{1+\theta r\sigma^2}$, implying that an altruistic principal offers a lower bonus than a selfish principal. This can also be seen after rewriting the first-order condition:

$$b_h = \frac{\delta + (1-\delta)\left[1-\gamma_h(\alpha_h)\right]^2}{\delta + (1-\delta)\left[1-\gamma_h(\alpha_h)\right]^2 + \theta r\sigma^2}.$$

The finding that $b_h < b_l$ implies that an altruistic principal pays a higher base salary than an egoistic principal. Recall the assumption made earlier that $r\sigma^2\theta \leq 1$, implying that PCL requires that a reduction in the bonus is compensated by a higher base salary. Thus, as long as PCL is satisfied, an altruistic principal always pays a higher base salary than an egoistic principal. The analysis thus leads to the following proposition:

Proposition 3 *In a separating equilibrium where the principal does not screen workers, an altruistic principal offers weaker financial incentives and pays a higher base salary than an egoistic principal.*

Given that the altruistic principal chooses the bonus that maximizes the expected surplus, how does he ensure that an egoistic principal is not willing to imitate him? This is the role of the base salary: the altruist increases the base salary up to the point that the selfish principal is not any longer willing to mimic him.

However, this does not imply that the altruistic principal pays a higher expected total compensation than his egoistic counterpart. The altruistic principal only pays a higher expected total compensation when workers provide more effort on average. To see this, it is convenient to use the fact that ICC1 is binding in equilibrium. Rewriting ICC1 gives:

$$b_h e_{\delta h} + s_h = b_l e_{\delta l} + s_l + e_{\delta h} - e_{\delta l}$$

Clearly, whether expected total compensation paid by the altruistic principal exceeds that of the selfish principal depends on the difference in average effort $e_{\delta h} - e_{\delta l}$. When the altruistic principal's equilibrium contract does not induce workers to provide more effort on average, credibly signaling altruism does not require paying a higher expected total compensation. As shown in the observable types case, conditionally altruistic workers do not always exert more effort when employed by an altruistic principal, because an altruistic principal sets weaker financial incentives, i.e. $b_h < b_l$. Therefore, we cannot be sure that a reciprocal worker exerts more effort when employed by an altruistic principal, whereas a selfish worker unambiguously provides less effort. Whether the altruistic principal pays a higher expected total compensation than an egoistic principal thus depends on the parameters.

Proposition 4 *The altruistic principal pays a higher expected total compensation than the egoist if and only if* $\gamma_h(\alpha_h) + r\theta\sigma^2 > 1$.

Proof. See appendix. ■

The intuition behind this condition is that the difference in average effort $e_{\delta h} - e_{\delta l}$ is only positive when the productivity of a conditionally altruistic worker sufficiently exceeds the productivity of an egoistic worker, either because a conditionally altruistic worker cares a lot about the principal's welfare or because it is costly to motivate the worker via financial incentives. So, combining propositions 3 and 4, paying a relatively low expected total compensation does not necessarily disprove a principal's altruism as long as it is accompanied by weak financial incentives.

The equilibrium is illustrated by figure 1. The figure shows ICC1 for $\gamma_h(\alpha_h) < \frac{1}{2}$ and the participation constraints PCL and PCH. Thus, ICC1 represents an isoprofit curve that for each bonus indicates the minimum base salary required to keep the selfish principal from imitating.[14] Similarly, PCL and PCH are indifference curves indicating the lowest base salary that is

[14]Specifically, ICC1 represents the isoprofit curve of the egoistic principal who imitates the altruist. This is identical to the material part of the altruistic principal's isopayoff curve.

acceptable to a selfish and reciprocal worker, respectively. The arrows thus demonstrate the area of feasible contracts. The dotted line represents an indifference curve of a hypothetical 'average' worker. That is, it is a weighted average of both worker types' indifference curves representing the utility levels they obtain in the optimum. The optimum is where the indifference curve of the hypothetical 'average' worker is tangent to ICC1. The corresponding bonus (b_h) maximizes the expected surplus.

For the remainder of the paper, it is important to understand the intuition behind the curves. The slope of ICC1 depends on $\gamma_h(\alpha_h)$:

Lemma 4 *Let s_h^{ICC1} denote the base salary that keeps the egoist from imitating, as defined by ICC1. s_h^{ICC1} is always decreasing in the bonus provided $(1-\delta)\gamma_h(\alpha_h) > \frac{1}{2}$, but has an inverted u-shape when $(1-\delta)\gamma_h(\alpha_h) < \frac{1}{2}$.*

Proof. See appendix. ■

The intuition is that ICC1 consists of two effects. On the one hand, worker's effort is increasing in the bonus, requiring an increase in the base salary to keep the selfish principal from mimicking. On the other hand, an increase in the bonus reduces the share of the marginal product that accrues to the principal, allowing for a decrease in the base salary. Since highly reciprocal workers exert relatively high effort and are relatively insensitive to incentive pay, the latter effect dominates when $(1-\delta)\gamma_h(\alpha_h) > \frac{1}{2}$.

Both PCL and PCH slope downwards, as depicted in figure 1. The reason is that an increase in the expected bonus payment benefits the worker, implying that the base salary should decrease to keep expected utility constant. This effect is partially offset by an increase in the worker's exposure to risk.[15] The reciprocal worker's PC, PCH, is always below PCL, which follows from the fact that a reciprocal worker derives more utility from the same equilibrium contract than an egoist. Moreover, PCH has a steeper slope than PCL:

Lemma 5 *Let s_h^{PCL} and s_h^{PCH} denote the base salary such that PCL and PCH are exactly satisfied. s_h^{PCL} and s_h^{PCH} have the following properties:*

$$\frac{ds_h^{PCH}}{db_h} < \frac{ds_h^{PCL}}{db_h} < 0.$$

Proof. See appendix. ■

The reason is that a conditionally altruistic worker exerts more effort than an egoist for a given bonus, implying that an increase in the bonus leads to

[15] This is by assumption, as I imposed that $r\sigma^2\theta \leq 1$ to ensure that indifference curves are downward sloping.

a larger increase in expected payments for a conditionally altruistic worker than for an egoistic worker. Therefore, a reciprocal worker permits a larger decrease in the base salary while keeping expected utility at the same level.

Two important observations need to be made. The first is that there is always a point on PCL that represents the contract offered by the selfish principal, namely (b_l, s_l). Since PCL is the selfish worker's indifference curve yielding his reservation utility $E(u_l) = \overline{u}$, the contract (b_l, s_l) is necessarily a point on PCL, as depicted in figures 1 and 2. The second observation is that when the altruistic principal offers b_h equal to b_l, ICC1 requires that $s_h > s_l$. The reason is that all contracts on ICC1 are assumed to succeed in signaling (and screening in the next section). Since for a given bonus expected effort is higher when the principal is believed to be an altruist, it must be that for $b_h = b_l$, $s_h > s_l$ to discourage the egoistic principal from imitating.

Finally, recall that I still have to show that the equilibrium contract (b_h, s_h) satisfies PCL. As argued above, the altruistic principal maximizes the utility of a hypothetical 'average worker', as workers are egoistic with probability δ. By lemma 5, the indifference curves of the 'average' worker are steeper than those of the egoist. Suppose that the altruistic principal chooses $b_h = b_l$ and $s_h > s_l$ to discourage the egoistic principal from imitating. This contract ensures that the egoistic worker's participation constraint is satisfied. Suppose that the altruistic principal chooses a contract (b_h, s_h), where $b_h < b_l$, on the 'average' worker's indifference curve that intersects this contract. As the average indifference curve is steeper than PCL, the selfish worker gains from a decrease in the bonus $(b_h < b_l)$ and a corresponding raise in the base salary that keeps the utility of the 'average' worker unchanged. Thus, the equilibrium contract ensures that the egoistic worker's participation constraint is satisfied.

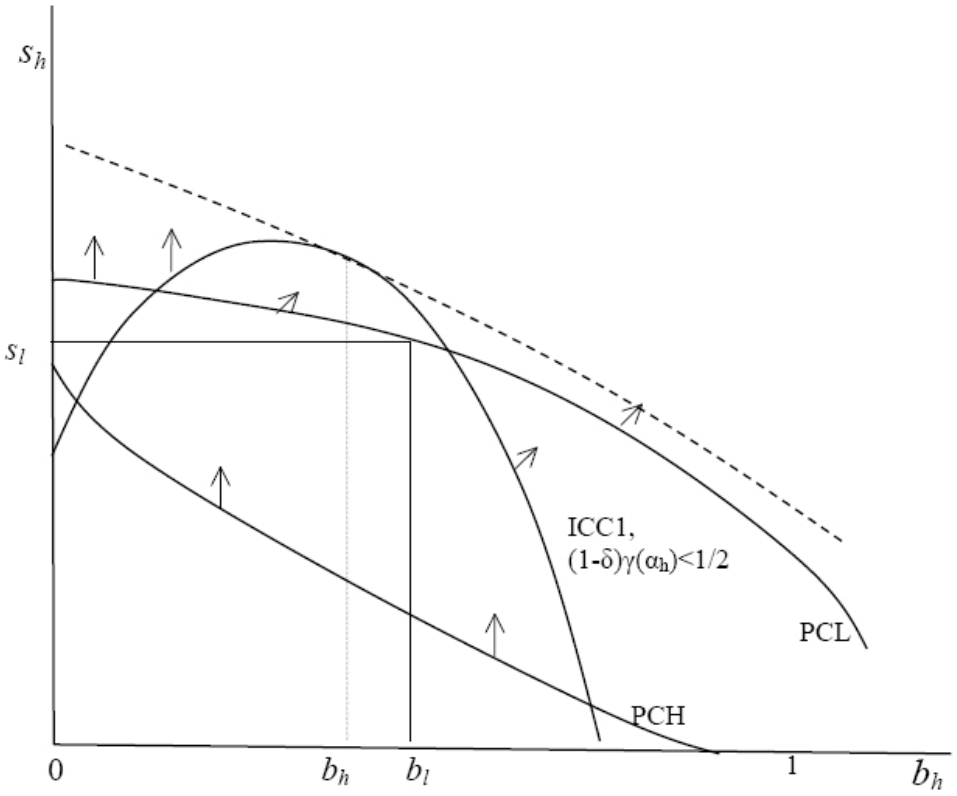

Figure 1: separating equilibrium without screening

5.3.2 Separating equilibrium: screening

This section studies a separating equilibrium where the altruistic principal writes a contract that signals his altruism and simultaneously screens worker types. First, consider the egoistic principal's contract choice. Because all workers behave completely egoistic when they believe that the principal is selfish (i.e. $\gamma_i(\alpha_l) = 0$), a selfish principal has no signaling or screening motive in equilibrium. Thus, as derived formally in lemma 1, the selfish principal offers a bonus $b_l = \frac{1}{1+r\sigma^2\theta}$ and the lowest base salary workers are willing to accept.

The altruistic principal may benefit from writing a contract that selects reciprocal workers only. As noted earlier, the altruistic principal's utility is increasing in the fraction of reciprocal workers, provided he convinces them that he is an altruist. Screening of worker types requires that the contract (b_h, s_h) simultaneously violates PCL and satisfies PCH. For ease of exposi-

tion, I refer to violating PCL as satisfying the screening constraint (SCC):

$$b_h e_{lh} + s_h - \frac{1}{2}\theta e_{lh}^2 - \frac{1}{2} r\sigma^2 b_h^2 \leq \overline{u}. \qquad \text{(SCC)}$$

The proposed equilibrium contract should not only screen worker types, but also signal the principal's altruism. Thus, the contract should satisfy the two incentive compatibility constraints. Assuming that SCC and PCH are satisfied, the incentive compatibility constraints ICC1 and ICC2 can be written as:

$$(1 - b_l)e_{\delta l} - s_l \geq (1 - b_h)e_{hh} - s_h \qquad \text{(ICC1')}$$

$$(1 - b_l)e_{\delta l} - s_l + \alpha_h \overline{u} \leq (1 - b_h)e_{hh} - s_h + \alpha_h E(u_h) \qquad \text{(ICC2')}$$

Since ICC1' and ICC2' are nothing but special cases of ICC1 and ICC2, the same reasoning applies to show that ICC1' is always binding and hence ICC2' is slack. Because profits are increasing in the worker's altruism, a selfish principal is always willing to imitate an altruistic principal, unless the altruistic principal explicitly takes ICC1' into account, see lemma 3 for a formal proof. Moreover, we observed in the previous section that when the principal abstains from screening workers, PCL is always satisfied. Therefore, succesful screening requires that the principal takes SCC into account, in other words SCC is binding as well. This also proves that the reciprocal worker's PC is satisfied, because when SCC is binding, PCH is slack. Thus, the altruistic principal's contract offer can be characterized as follows:

Lemma 6 *In a separating and screening equilibrium, the altruistic principal offers a contract such that ICC1' and SCC hold with equality. Such a contract always exists for* $b_h \in (b_l, 1)$.

Proof. See appendix. ■

Figure 2 illustrates the altruistic principal's maximization problem. The similarity with figure 1 should be clear: it shows ICC1', SCC and PCH. Recall that the altruistic principal chooses the point on ICC1' that maximizes the reciprocal worker's expected utility. By shifting the conditionally altruistic worker's indifference curve (PCH) up along ICC1', it can easily be seen that, given the screening constraint, his expected utility is maximized at an intersection of ICC1' and SCC.

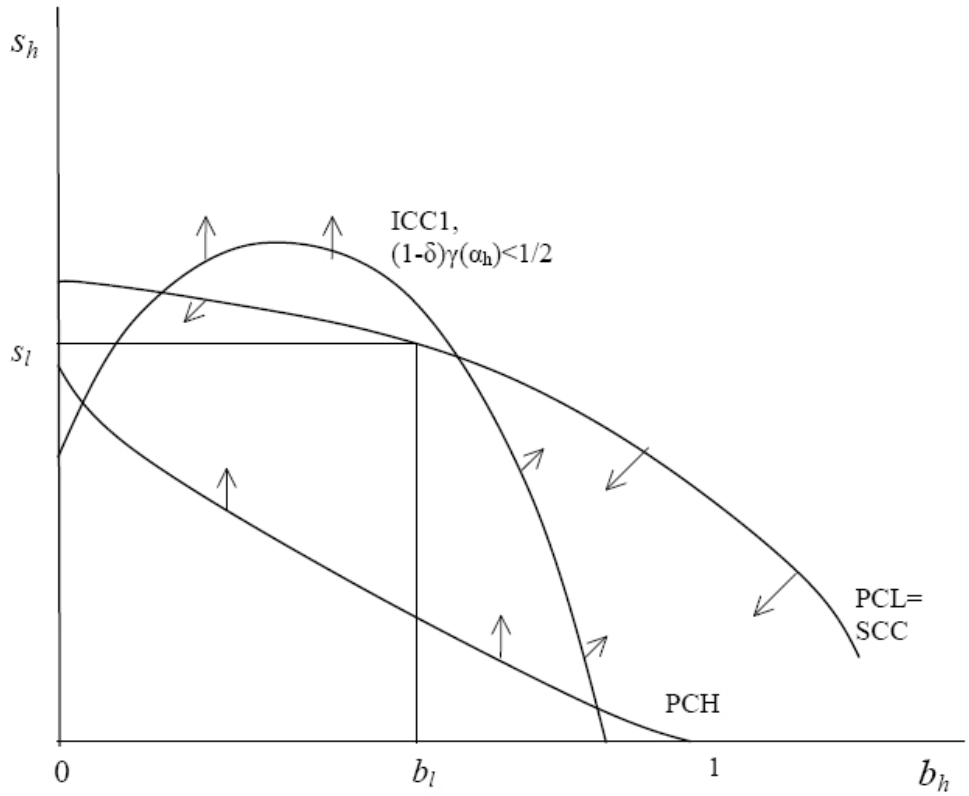

Figure 2: separating equilibrium where the altruistic principal screens

Figure 2 provides an intuitive explanation of lemma 6. Recall that for $b_h = b_l$, ICC1' lies above SCC. When $b_h = 1$, ICC1' is always below SCC, and hence an intersection point on the interval $(b_l, 1)$ always exists. There is a clear intuition for this fact. When $b_h = 1$, the worker is the full residual claimant and SCC thus specifies the maximum amount he is willing to pay for the firm. This amount is equal to the expected revenues minus the costs of effort, risk-bearing and the outside option. Similarly, when $b_h = 1$, ICC1' specifies the maximum amount an altruistic principal can receive for the firm: this amount should not exceed the profits made by the selfish principal. The selfish principal's profits are given by the expected revenues minus the compensation for the worker's effort, risk and outside option. The amount the worker is willing to pay for the firm (SCC) is always smaller than the equilibrium profits made by the selfish principal (ICC1'), because the selfish principal sets the bonus at the surplus-maximizing level ($b_l < 1$). By contrast, when the altruistic principal makes the worker full residual claimant ($b_h = 1$), he exposes the worker to an inefficient amount of risk, which reduces his willingness to pay for ownership of the firm. When expressed as a base

salary, SCC thus always exceeds ICC1', implying that an intersection point on the interval $(b_l, 1)$ always exists.

However, as illustrated by figure 2, when $\gamma_h(\alpha_h)$ is small, ICC1' has an inverted u-shape, implying that there are two intersection points.[16] In that case, the principal prefers the bonus at the intersection point on the interval $(b_l, 1)$. To see this, recall that the altruistic principal chooses the point on ICC1' that maximizes the reciprocal worker's expected utility. By shifting the conditionally altruistic worker's indifference curve (PCH) up along ICC1', it can easily be seen that, given the constraints, the principal prefers the intersection point that specifies the highest bonus. The reason is that an increase in the bonus is more beneficial for a reciprocal worker than an egoistic worker due to the former's higher effort.

Proposition 5 *In a separating and screening equilibrium, the altruistic principal offers stronger incentives and pays a lower base salary than the egoistic principal, i.e.* $b_h \in (b_l, 1)$ *and* $s_h < s_l$.

Proof. See appendix. ■

In contrast to the previous section, the altruistic principal pays a lower base salary than the egoistic principal. As illustrated by figure 2, the downward sloping participation constraint of the egoistic worker implies that, as this constraint is binding, a higher bonus is reflected in a lower base salary. Nevertheless, the expected total compensation earned/paid by the altruistic worker/principal is always larger than that of the selfish type.[17] The reason is that because a reciprocal worker faces stronger incentives than a selfish worker, he unambiguously provides more effort ($e_{hh} > e_{\delta l}$), implying that an altruistic principal has to pay more than a selfish principal to discourage him from imitating. This result is summarized in the following proposition:

Proposition 6 *In a separating and screening equilibrium, an altruistic principal pays a higher expected total compensation than an egoistic principal.*

These results stand in remarkable contrast with the results in the previous section. The reason for these diverging findings is that screening can (most

[16] A second intersection point may also exist when $\gamma_h(\alpha_h)$ is sufficiently high and $b_h > 1$. I assume that $b \leq 1$, but it can be shown that the intersection point on the interval $(b_l, 1)$ is strictly preferred.

[17] As in the previous section, this can be seen from rewriting ICC1':

$$b_h e_{hh} + s_h = b_l e_{\delta l} + s_l + e_{hh} - e_{\delta l},$$

implying that $b_h e_{hh} + s_h > b_l e_{\delta l} + s_l$, since $e_{hh} > e_{\delta l}$.

efficiently) be accomplished by offering stronger incentives than otherwise optimal. These excessively strong incentives reduce the total surplus and diminish the attractiveness of the contract for the selfish worker, which is inevitable in order to satisfy both the screening constraint and ICC1'.

5.3.3 Comparison of equilibria

One may wonder whether the altruistic principal prefers the separating and screening equilibrium above the standard separating equilibrium. This question is particularly relevant, because the altruistic principal can always deviate to the equilibrium that gives him the largest total expected payoff. As argued above, as long as ICC1 or ICC1' are satisfied, the intuitive criterion implies that deviation must come from an altruistic principal. Thus the equilibrium that yields the altruistic principal the highest payoff will be the unique equilibrium outcome.

Because the altruistic principal's profits are identical in the two situations (namely the same as the selfish principal's payoff $E(\pi_l^*)$), the equilibrium that yields the highest expected worker utility is preferred. Screening of worker types has the advantage that only conditionally altruistic workers are attracted, which has a positive effect on expected worker utility. First, because conditionally altruistic workers exert more effort, which is reflected in a higher salary. Second, because conditionally altruistic workers derive utility from the principal's welfare. However, screening is also costly: the bonus is distorted compared to the efficient bonus level. The higher $\gamma_h(\alpha_h)$, the larger the distortion. Thus, incurring the costs of screening is unattractive when $\gamma_h(\alpha_h)$ is relatively large and when the large majority of workers is conditionally altruistic:

Proposition 7 *A separating and screening equilibrium always exists for small values of $\gamma_h(\alpha_h)$. Screening is less likely, the higher $\gamma_h(\alpha_h)$ and the smaller δ.*

Proof. See appendix. ■

6 Concluding remarks

I have studied the relation between monetary gift-exchange and incentives by incorporating reciprocity in an otherwise standard principal-agent model. The specification of reciprocity is taken from Levine (1998), and allows to distinguish between authentic and strategic kindness. The key assumption is that some workers care more for the principal when they are convinced

that the principal cares for them. The principal can be egoistic or altruistic. An altruistic principal can signal his type by offering a generous contract, consisting of a base salary and a piece rate. As is common in principal-agent models, the worker is risk-averse and not constrained by limited-liability. Inspired by the findings in several experiments, I have allowed for worker heterogeneity by assuming that not all workers are reciprocal. As a result, the principal may find it attractive to screen workers.

Assuming that types are private information, I have found that an altruistic principal who abstains from screening, signals his altruism by offering relatively weak incentives and a relatively high base salary. The piece rate and the base salary simultaneously convince the worker of the principal's care. The reason for offering weak incentives is that when workers are convinced of the principal's care, strong incentives add little to worker's productivity, while exposing workers to unnecessary risk. Offering strong incentives is therefore suboptimal for a principal who cares about the worker's well-being.

The second finding is that to induce reciprocity, an altruistic principal does not necessarily have to pay a higher expected total compensation than an egoistic principal. As part of the altruistic principal's gift is a reduction in incentives, the worker's expected effort may be relatively low, despite his altruistic feelings. Therefore, an egoistic principal has no reason to mimic the altruist, implying that a relatively low expected total compensation may suffice to distinguish both types. An altruistic principal only pays a higher expected total compensation when signaling altruism leads to significant productivity gains.

Finally, as some workers do not reciprocate the principal's altruism, the principal may find it optimal to write a contract that simultaneously signals his altruism and screens reciprocal worker types. I have shown that such a contract is characterised by excessively strong incentives and a relatively high expected total compensation. Incentives are a suitable instrument for screening workers, because conditionally altruistic workers put in more effort than egoistic workers and hence gain more from output-contingent pay. Thus, strong incentives are offered to attract the worker who needs them the least.

Of course, there are some limitations to the analysis. A first limitation is that I only looked at monetary rewards, while employers typically have other instruments to stimulate or control workers, such as work rules, work organisation, minimum effort requirements, task assignment or giving personal attention to workers. All of these may be helpful to signal the employer's benevolent intentions. Therefore, it may well be that employers that establish good relationships with their workers via these non-monetary means can afford paying lower wages, as in Dur (2009) and Dur et al. (2010). This would be well in line with the results of the observable types case.

Relatedly, a second limitation is that in large organizations wages are not always determined by the relevant managers, and the interpretation as a gift may therefore be problematic. It would be interesting to see how wage-setting institutions and organizational structure impact on the prospects for gift-exchange.

A third limitation is that the model I presented is a partial equilibrium model: there is no competition between employers and the outside option of both worker types is exogenously given. It is not ex-ante clear to what extent the results carry over to a general equilibrium setting. Thus, there is ample room for further research, both theoretically and empirically.

7 Appendix

7.1 Proof of proposition 2

The proof builds on the analysis of the separating equilibria. Screening requires that the egoistic worker's participation constraint (PCL) is violated, while at the same time the conditionally altruistic worker's participation constraint is satisfied. All contracts that satisfy these two conditions are possible equilibrium candidates, depending on the out-of-equilibrium beliefs. I show that none of these contracts, denoted (b_p, s_p), is an equilibrium contract, as there is always a profitable deviation.

First, we can rule out all pooling contracts that yield the egoistic principal lower expected profits than he earns in a separating equilibrium, $E(\pi_l^*)$. The egoistic principal can always profitably deviate by offering the separating equilibrium contract (b_l, s_l), which ensures him at least $E(\pi_l^*)$, irrespective of the worker's out-of-equilibrium beliefs.

To rule out all pooling contracts (b_p, s_p) that yield higher profits than $E(\pi_l^*)$, we proceed along the lines suggested in the main text. We look for an alternative contract offer (b', s') that can be profitable for an altruistic principal, but not for an egoist. The reason is that for such a deviation, the Intuitive Criterion uniquely defines the out-of-equilibrium beliefs, namely that $pr(\alpha_j = \alpha_h) = 1$. More formally, the contract (b', s') breaks the equilibrium if the following conditions are satisfied:

1. The expected monetary payoff with the most optimistic beliefs possible is equal to the monetary payoff in equilibrium.

2. A conditionally altruistic worker obtains higher expected utility than when the equilibrium contract (b_p, s_p) is offered.

3. The screening contraint SCC is satisfied, i.e. the egoistic worker's participation constraint PCL is violated.

That such a contract always exists can be shown as follows. Let ICC1" represent the same isoprofit curve as ICC1', but for a higher level of profits, namely $E(\pi_p) > E(\pi_l^*)$. Mathematically:

$$(1 - b')e_{hh} - s' = E(\pi_p), \qquad \text{(ICC1")}$$

Thus, ICC1" represents all contracts (b', s') that yield the same profit level $E(\pi_p)$ as the equilibrium contract (b_p, s_p), assuming the worker is conditionally altruistic and that he believes that he is employed by an altruistic principal. As higher profit levels are represented by lower indifference curves, all pooling contracts that yield higher monetary profits than (b_l, s_l) have a corresponding isoprofit curve ICC1" below ICC1'.

Suppose the altruistic principal increases the base salary up to the point that monetary profits equal equilibrium profits, i.e. $b' = b_p$, $s' > s_p$ such that ICC1" holds. There are two cases. First, the screening constraint SCC is satisfied. Clearly, such a contract meets all of the conditions for profitable deviation. Second, the screening constraint SCC is violated. In that case the isoprofit curve ICC1" always contains another contract that meets the conditions outlined above, characterised by $b' \in (b_p, 1)$. As ICC1" is nothing but ICC1' for a higher profit level, it follows from the properties of ICC1' that for $b' = b_p$, $s' > s_p$, while for $b' = 1$ PCL is violated (or SCC satisfied). Thus, in this case there always exists a contract $b' \in (b_p, 1)$ and corresponding base salary s' such that PCL and ICC1" are exactly satisfied. This contract therefore also yields a conditionally altruistic worker higher utility: increasing the base salary until PCL is satisfied, while keeping b constant, gives a conditionally altruistic worker a higher utility than in equilibrium. As the participation constraint of a conditionally altruistic worker is steeper than PCL, offering $b' \in (b_p, 1)$ further increases his utility.

7.2 Proof of proposition 4

As ICC1 is binding, the altruistic principal pays a higher expected total compensation if and only if $e_{\delta h} > e_{\delta l}$. Inserting expressions for effort, b_l and b_h into this condition, we obtain after considerable rewriting:

$$\frac{(1-\delta)\, r\sigma^2 \gamma_h(\alpha_h)}{(1+\theta r\sigma^2)} \frac{\gamma_h(\alpha_h) + r\theta\sigma^2 - 1}{\delta + (1-\delta)\left[1-\gamma_h(\alpha_h)\right]^2 + \theta r\sigma^2} > 0,$$

implying that the altruist pays more if and only if

$$\gamma_h(\alpha_h) + r\theta\sigma^2 > 1.$$

7.3 Proof of lemma 4

It follows from straightforward rewriting of ICC1 that s_h^{ICC1} is defined as

$$s_h^{ICC1} = (1 - b_h)e_{\delta h} - (1 - b_l)e_{\delta l} - s_l,$$

or rewritten, using equations (5) and (6):

$$s_h^{ICC1} = (1 - b_h)\left(\frac{b_h}{\theta} + (1 - \delta)\,\frac{\gamma_h(\alpha_h)(1 - b_h)}{\theta}\right) - \frac{1}{2\theta}\frac{1}{(1 + \theta r\sigma^2)} + \overline{u}.$$

Inspection of the derivative with respect to b_h proves that s_h^{ICC1} initially increases in the bonus provided $(1 - \delta)\,\gamma_h(\alpha_h) < \frac{1}{2}$, but always decreases in the bonus when $(1 - \delta)\,\gamma_h(\alpha_h) > \frac{1}{2}$:

$$\frac{ds_h^{ICC1}}{db_h} = \frac{-2\left[b_h + (1 - \delta)\,\gamma_h(\alpha_h)(1 - b_h)\right] + 1}{\theta}.$$

7.4 Proof of lemma 5

Assuming that PCL or, equivalently, SCC holds with equality, PCL can be rewritten to:

$$s_h^{PCL} = \overline{u} - \frac{b_h^2}{2\theta}\left(1 - \theta r\sigma^2\right).$$

Clearly, s_h^{PCL} decreases in the bonus by assumption, as we imposed that $\theta r\sigma^2 < 1$ to prevent un upward slope. For ease of comparison, I provide the derivative with respect to b_h:

$$\frac{ds_h^{PCL}}{db_h} = -\frac{b_h}{\theta}\left(1 - \theta r\sigma^2\right).$$

Similarly, s_h^{PCH} follows from rewriting PCH:

$$s_h^{PCH} = \overline{u} - \frac{b_h^2}{2\theta}\left(1 - \theta r\sigma^2\right) + \frac{\gamma_h(\alpha_h)^2}{2\theta}\left(1 - b_h\right)^2 - \gamma_h(\alpha_h)E(\pi_h).$$

Keeping $E(\pi_h)$ constant at its equilibrium level $E\left(\pi_l^*\right)$ and taking the derivative with respect to b_h, it can easily be seen that compared to PCL, PCH is steeper as long as $b_h \leq 1$:

$$\frac{ds_h^{PCH}}{db_h} = -\frac{b_h}{\theta}\left[1 - \theta r\sigma^2\right] - \frac{\gamma_h(\alpha_h)^2}{\theta}\left[1 - b_h\right].$$

7.5 Proof of lemma 6

Because both ICC1' and SCC are binding, the altruistic principal maximizes his payoff by offering a contract such that both constraints hold with equality. To show that such a contract always exists, first note that for $b_h = b_l$, ICC1' requires a higher base salary s_h than SCC. Since for a given bonus, a conditionally altruistic worker exerts more effort when he believes that the principal is an altruist, signaling is only credible if $s_h > s_l$. It can be shown that when $b_h = 1$, ICC1' always allows for a lower base salary than SCC, and hence an intersection point on the interval $(b_l, 1)$ always exists.

Recall that s_h^{PCL} denotes the maximum salary SCC allows for, see lemma 5. Inserting $b_h = 1$ into SCC, $-s_h^{PCL}$ describes a selfish worker's willingness to pay for the firm:

$$-s_h^{PCL} = e_{lh} - \frac{1}{2}\theta e_{lh}^2 - \frac{1}{2}r\sigma^2 - \overline{u}.$$

Recall that s_h^{ICC1} denotes the base salary that keeps the egoistic principal from imitating, see lemma 4. As ICC1' is nothing but a special case of ICC1, $s_h^{ICC1'}$ can be obtained by inserting $\delta = 0$ into s_h^{ICC1}. Inserting $b_h = 1$ into ICC1', we obtain that $-s_h^{ICC1'} = E(\pi_l^*)$. Thus, the maximum amount the altruistic principal can receive for the firm $(-s_h^{ICC1'})$ is equal to the egoistic principal's expected equilibrium profits. Using equations (8) and (10), we can write $E(\pi_l^*)$ as:

$$-s_h^{ICC1'} = E(\pi_l^*) = e_{ll} - \frac{1}{2}\theta e_{ll}^2 - \frac{1}{2}r\sigma^2 b_l^2 - \overline{u}.$$

Since b_l is the surplus maximizing bonus when workers are egoistic, $E(\pi_l^*) > -s_h^{PCL}$, implying that $s_h^{ICC1'} < s_h^{PCL}$ when $b_h = 1$.

7.6 Proof of proposition 5

By lemma 6, the equilibrium contract satisfies ICC1' and SCC with equality, and such a contract always exists for $b_h \in (b_l, 1)$. The altruistic principal prefers this contract over possible contracts that also satisfy ICC1' and SCC with equality for $b_h < b_l$. As the principal's monetary payoff is constrained by ICC1', he chooses the contract that maximizes a conditionally altruistic worker's utility. By lemma 5, a conditionally altruistic worker's indifference curve has a steeper slope than an egoistic worker's indifference curve. Hence, utility of a conditionally altruistic worker is maximized when the principal chooses $b_h \in (b_l, 1)$ rather than $b_h < b_l$. The downward sloping indifference curves imply that $s_h < s_l$.

7.7 Proof of proposition 7

To show under what conditions the altruistic principal prefers to screen, first note that when $\gamma_h(\alpha_h) = 0$, the payoff is equal in both equilibria. I now analyze how the principal's expected payoff changes when $\gamma_h(\alpha_h)$ increases in both equilibria, where the change in the principal's expected payoff is equal to the change in the worker's expected utility.

In a separating equilibrium where the principal does not screen, the effect of a change in $\gamma_h(\alpha_h)$ on total utility is given by $\delta \frac{dE(u_l)}{d\gamma_h(\alpha_h)} + (1-\delta)\frac{dE(u_h)}{d\gamma_h(\alpha_h)}$, or after some rewriting:

$$\delta \frac{dE(u_l)}{d\gamma_h(\alpha_h)} + (1-\delta)\frac{dE(u_h)}{d\gamma_h(\alpha_h)} = (1-\delta)\left[(1-b_h)\frac{de_{hh}}{d\gamma_h(\alpha_h)} + E(\pi_h)\right].$$

This has a simple interpretation; the gain in total utility when $\gamma_h(\alpha_h)$ increases is equal to the additional productivity of the reciprocal type (reflected in the base salary), plus his increased utility from the immaterial aspect of the job. Worker's utility is convex in $\gamma_h(\alpha_h)$, because b_h is decreasing and hence $\frac{de_{hh}}{d\gamma_h(\alpha_h)}$ is increasing in $\gamma_h(\alpha_h)$.

Similarly, in a separating and screening equilibrium the effect of a change in $\gamma_h(\alpha_h)$ on total expected utility is given by $\frac{dE(u_h)}{d\gamma_h(\alpha_h)}$, or:

$$\frac{dE(u_h)}{d\gamma_h(\alpha_h)} = \frac{db_h}{d\gamma_h(\alpha_h)}\left[\frac{de_{hh}}{db_h}(1-\theta e_{hh}) - r\sigma^2 b_h\right] + (1-b_h)\frac{de_{hh}}{d\gamma_h(\alpha_h)} + E(\pi_h).$$

The first part is negative and represents the loss in worker utility because b_h is suboptimally high. The second part has a similar interpretation as in a signaling equilibrium, but keeping the bonus constant, it is larger because the worker is reciprocal with probability 1 instead of $(1-\delta)$. The first part is zero if $\gamma_h(\alpha_h) = 0$, but becomes smaller (increases in absolute value) when $\gamma_h(\alpha_h)$ becomes larger. The second part is concave in $\gamma_h(\alpha_h)$, because b_h is increasing and hence $\frac{de_{hh}}{d\gamma_h(\alpha_h)}$ is decreasing in $\gamma_h(\alpha_h)$.

Comparing $\delta \frac{dE(u_l)}{d\gamma_h(\alpha_h)} + (1-\delta)\frac{dE(u_h)}{d\gamma_h(\alpha_h)}$ and $\frac{dE(u_h)}{d\gamma_h(\alpha_h)}$, the separating and screening equilibrium is always preferred for $\gamma_h(\alpha_h)$ sufficiently close to zero. When $\gamma_h(\alpha_h)$ increases, the cost of distorting the bonus becomes more severe, and at some point the standard separating equilibrium will be preferred. The larger the fraction of selfish workers δ, the higher the values of $\gamma_h(\alpha_h)$ that sustain a screening equilibrium.

References

[1] Akerlof, George A. (1982), "Labor Contracts as Partial Gift Exchange." *Quarterly Journal of Economics*, 97(4): 543-569.

[2] Arbak, Emrah and Laurence Kranich (2005), "Can Wages Signal Kindness?" GATE Working Paper 05-11.

[3] Barro, Robert (1974): "Are Government Bonds Net Wealth?", *Journal of Political Economy*, 82(6): 1095-1117.

[4] Bartling, Björn, Ernst Fehr, and Klaus M Schmidt (2011), "Screening, Competition, and Job Design: Economic Origins of Good Jobs" *American Economic Review*, forthcoming.

[5] Battigalli, Pierpaolo and Martin Dufwenberg (2009), "Dynamic Psychological Games." *Journal of Economic Theory*, 144(1): 1-35.

[6] Becker, Gary (1974): "A Theory of Social Interactions." *Journal of Political Economy*, 82(6): 1063-1094.

[7] Bellemare, Charles and Bruce Shearer (2011), "On the Relevance and Composition of Gifts within the Firm: Evidence from Field Experiments." *International Economic Review*, forthcoming

[8] Bernheim, B.D. and Oded Stark (1988): "Altruism within the Family Reconsidered: Do Nice Guys Finish Last?", *American Economic Review*, 78(5), 1034-1045.

[9] Bewley, Truman F. (1999), *Why Wages Don't Fall during a Recession*. Harvard University Press.

[10] Charness, Gary and David Levine (2007), "Intention and Stochastic Outcomes: An Experimental Study." *Economic Journal*, 117(522): 1051-1072.

[11] Cho, In-Koo and David M. Kreps (1987), "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics,* 102(2): 179-221.

[12] Cropanzano, Russel and Marie S. Mitchell (2005), "Social Exchange Theory: An Interdisciplinary Review." *Journal of Management*, 31(6): 874-900.

[13] Dufwenberg, Martin and Georg Kirchsteiger (2004), "A Theory of Sequential Reciprocity." *Games and Economic Behavior*, 47(2): 268-298.

[14] Dur, Robert (2009), "Gift Exchange in the Workplace: Money or Attention?" *Journal of the European Economic Association*, 07(2-3): 550-560.

[15] Dur, Robert, Arjan Non, and Hein Roelfsema (2010), "Reciprocity and Incentive Pay in the Workplace." *Journal of Economic Psychology*, 31(4): 676-686.

[16] Dur, Robert and Joeri Sol (2010): "Social Interaction, Co-Workers Altruism, and Incentives." *Games and Economic Behavior*, Vol. 69(2): 293-301.

[17] Englmaier, Florian and Stephen Leider (2008), "Contractual and Organizational Structure with Reciprocal Agents." CESifo Working Paper 2415.

[18] Falk, Armin and Urs Fichbacher (2006), "A Theory of Reciprocity." *Games and Economic Behavior*, 54(2): 293-315.

[19] Falk, Armin and Michael Kosfeld (2006), "The Hidden Costs of Control." *American Economic Review*, 96(5), 1611–30.

[20] Fehr, Ernst and Simon Gächter (2000), "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives*, 14(3): 159-181.

[21] Fehr, Ernst and Simon Gächter (2002), "Do incentive contracts undermine voluntary cooperation?" Working Paper No. 34, Institute for Empirical Research in Economics, University of Zürich.

[22] Fehr, Ernst, George Kirchsteiger and Arno Riedl (1993), "Does Fairness Prevent Market Clearing? An Experimental Investigation." *Quarterly Journal of Economics*, 108(2): 437-459.

[23] Fehr, Ernst and John A. List (2004), "The Hidden Costs and Returns of Incentives – Trust and Trustworthiness Among CEOs." *Journal of the European Economic Association*, 2(5): 743-771.

[24] Fischbacher, Urs, Simon Gächter, and Ernst Fehr (2001) "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment." *Economics Letters*, 71(3): 397–404.

[25] Fischbacher, Urs, and Simon Gächter (2010) "Social Preferences, Beliefs, and the Dynamics of Free Riding in Public Goods Experiments." *American Economic Review*, 100(1): 541–56.

[26] Gächter, Simon and Christian Thöni (2010), "Social Comparison and Performance: Experimental Evidence on the Fair Wage-Effort Hypothesis." IZA Working paper No. 4687.

[27] Gul, Faruk, and Wolfgang Pesendorfer (2010), "Interdependent Preference Models as a Theory of Intentions." Princeton University Working Paper.

[28] Kosfeld, Michael and Ferdinand von Siemens (2009), "Worker Self-Selection and the Profits from Cooperation." *Journal of the European Economic Association,* 7(2-3): 573-582.

[29] Kosfeld, Michael and Ferdinand von Siemens (2011), "Competition, Cooperation, and Corporate Culture." *The RAND Journal of Economics*, 42(1): 23-43.

[30] Levine, David K. (1998), "Modeling Altruism and Spitefulness in Experiments." *Review of Economic Dynamics*, 1(3): 593-622.

[31] Ley, Eduardo (1997), "Optimal Provision of Public Goods with Altruistic Individuals." *Economic Letters*, 54(1), 23-27.

[32] Rabin, Matthew (1993), "Incorporating Fairness into Game Theory and Economics." *American Economic Review*, 83(5): 1281-1302.

[33] Rhoades, Linda, and Robert Eisenberger (2002), "Perceived Organizational Support: A Review." *Journal of Applied Psychology*, 87(4), 698-714.

[34] Sliwka, Dirk (2007), "Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes." *American Economic Review*, 97(3), 999-1012.

[35] Soberman, D., 2003. "Simultaneous Signaling and Screening with Warranties." *Journal of Marketing Research*, 40 (2): 176-209.

[36] Sol, Joeri, (2010), "Peer Evaluation: Incentives and Co-Worker Relations." Tinbergen Institute Discussion Paper 10-055/1.