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†

January 5, 2010

Abstract

Using a formal principal-agent model, I investigate the relation between monetary gift-exchange and incentive pay, while allowing for worker heterogeneity. I assume that some agents care more for their principal when they are convinced that the principal cares for them. Principals can signal their altruism by offering a generous contract, consisting of a base salary and an output-contingent bonus. I find that principals signal their altruism by offering relatively weak incentives and a relatively high expected total compensation, but the latter does not necessarily hold. Furthermore, since some agents 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 show that such a contract is characterised by excessively strong incentives and 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, Oleg Shchetinin, Joeri Sol, Paul Steffens and seminar participants in Rotterdam for many useful comments.

†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 employer and employee, where employee's effort and employer's benevolent treatment of workers are reciprocal gifts.

Benevolent treatment of employees encompasses several aspects. Of all the aspects mentioned by Akerlof, the wage level has without doubt attracted most attention. Numerous laboratory experiments find a positive relation between employee's effort and the salary offered by the employer.[1] However, the question how the wage level interacts with monetary incentives in establishing gift-exchange relationships has received far less attention, especially in theory but also in experiments. This is remarkable, since monetary incentives belong to the core of economics. Moreover, the lack of theoretical investigation is a significant issue, because theoretical models usually allow for more generalizable results than a specific experimental setting. In a typical experiment, the contracting choices are restricted, exposure to risk is very limited, and players are randomly matched into pairs by the experimenter. By contrast, in the real world employers are far less restricted in their contract choices, incentive schemes typically expose workers to risk, and employers may screen workers by their contract choices. Therefore, several subtle effects may come into play when the environment becomes increasingly realistic.

In this paper, I use a formal principal-agent model to study gift-exchange in a setting that captures some important complexities of the real world. The model is inspired by Levine (1998)'s game-theoretic approach of modeling reciprocity. The essence of this approach is that the intensity of an individual's altruism is conditional on his belief about the other individual's altruism, i.e. an individual cares more for someone who is perceived to care for him. Applied to the principal-agent model, I assume that the principal is either altruistic or selfish, and that whether agents are altruistic or selfish de-

[1] An early experimental study is Fehr, Kirchsteiger and Riedl (1993). Fehr and Gächter (2000) survey the voluminous literature.

pends on their beliefs about the principal's type. However, a typical finding in laboratory experiments is that not all individuals are equally motivated by reciprocal tendencies. Therefore, I assume that not all workers are reciprocal. Altruistic principals design a contract, consisting of a base salary and an output -contingent bonus, that signals their altruism and at the same time provides proper incentives to exert effort. Moreover, the contract can be designed to screen the reciprocal workers. Compared to existing gift-exchange models, there are thus three innovations. The key innovation is that the contractual environment allows for an unrestricted choice of base salary and bonus payments. Another innovation is that I examine the consequences of worker heterogeneity. In particular, I explore how contracts can be designed to ensure self-selection of reciprocal worker types and at the same time signal the principal's compassionate predisposition. Finally, I also investigate the situation where workers can tell from their experience in daily interactions with their boss whether he genuinely cares for them or not. Although an extreme case, the findings yield qualitative predictions for long-term work relationships.

The results reveal that monetary gift-exchange is a context-dependent phenomenon. First, I show that when principals abstain from screening workers, gift-exchange is associated with weaker incentives. When gift-exchange relations are established, there is less need to stimulate workers via monetary incentives. Second, I find that the positive relation between wages and effort may be obscured when principals are not restricted in their contract choices. Thus, although in our model a generous wage offer is always interpreted as a signal of altruism, offering a high expected total compensation is not a prerequisite to succesfully signal one's type. Because gift-exchange relations induce the altruistic principal to weaken monetary incentives, worker's effort is relatively low, and hence a relatively low total compensation sometimes suffices to distinguish an altruistic from a selfish principal. Third, and perhaps surprisingly, excessively strong incentives may result when the principal attempts to signal his benevolent intentions and simultaneously select motivated workers. The paradox is that strong incentives are offered in order to attract the employees who need them the least. Fourth, when the principal's care for his workers is self-evident, the positive relation between wages and effort vanishes. In fact, the high quality of the relation may even be reflected in a compensating wage differential. Moreover, a high quality of the relation weakens the attractiveness of offering strong monetary incentives.

The paper is closely related to an emerging experimental literature on incentives, contract design and reciprocity. The main findings of this literature are compatible with my model. An important issue is whether monetary incentives crowd out the possibilities for gift-exchange. Inspired by a substantial literature in social psychology, it has been suggested that material rewards may undermine the intrinsic motivation to perform a task, see e.g. Frey (1997). In the most radical interpretation, the introduction of incentives may thus end all voluntary cooperation, where voluntary cooperation is defined as the difference between actual and privately optimal effort. This complete crowding-out hypothesis is generally inconsistent with my model. A weaker interpretation is that incentives partially crowd out voluntary cooperation. Partial crowding out means that keeping the wage constant, voluntary cooperation is lower when (stronger) incentives are implemented. My model is in line with the partial crowding-out hypothesis. According to my model, a tight link between effort and reward restricts the worker's opportunities to reciprocate a principal's favourable treatment, and hence weakens the importance of the worker's altruism. The complete crowding-out hypothesis is hard to reconcile with the experimental evidence. A consistent finding is that the presence of incentives, when framed as a reward, does not preclude a positive relation between total compensation and effort, see Güth et al. (1998), Anderhub et al. (2002), Fehr and Gächter (2002), Bellemare and Shearer (2007), Gächter et al. (2009) and Rigdon (2009).[2] By contrast, Fehr and Gächter (2002) and Bellemare and Shearer (2009) present evidence in line with the partial-crowding out hypothesis: the positive relation between effort and wages is weaker when incentives are stronger. In other words, although incentives increase total effort, they reduce voluntary cooperation for a given wage level.

This paper contributes to this literature by pointing at some issues that have received little or no attention yet. First, my model shows that the crowding effect will be stronger, the more direct the link between effort and rewards, i.e. the larger the piece rate. In most of the experimental literature, however, the incentive scheme is non-linear: a bonus is paid or a fine imposed conditional on the effort level. It may be worthwhile to investigate the issue of partial crowding for alternative incentive structures as well. Second, our model shows that in experiments that provide the principal with multiple instruments, the method of matching participants may carry important im-

[2]Three remarks are worth making here. First, Güth et al. (1998) do not deal with the issue of crowding-out, but their results show that gift-exchange is not inhibited by incentives. Second, in Gächter et al. (2009), there is either no crowding or complete crowding depending on experiencing a trust session without incentives first. Finally, the study of Bellemare and Shearer (2007) is noteworthy, because it is the only field study that considers a firm where workers are paid according to piece rates and it is the only field study that finds strong evidence in favor of the monetary gift-exchange. Field experiments by Gneezy and List (2006), Kube et al. (2006, 2008), Al-Ubaydli et al. (2007) and Hennig-Schmidt et al. (2008) find only limited support for monetary gift-exchange.

plications. Specifically, a competitive labor market may induce principals to set stronger incentives than when agents are matched randomly.

The paper proceeds as follows. The next section describes the related literature. Then, in section 3 I set out the model and analyze the observable types case, which serves as a benchmark for the next two sections where types are assumed to be unobservable. In section 4.1, I introduce some simplifying assumptions in order to analyze in section 4.2 how monetary incentives and the base salary interact in establishing a gift-exchange relationship. In section 4.3 I study how contracts in addition can help to select reciprocal workers. Finally, in section 5 I conclude and provide some avenues for further research.

2 Related literature

In economics, several authors have suggested ways of modeling reciprocity, 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). Despite the recent attention for modeling reciprocal behavior in the workplace, there are only few studies that investigate the relation between gift-exchange and monetary incentives.

One of them is Englmaier and Leider (2008). They also present a principal-agent model with conditionally altruistic and risk averse agents, allowing for endogenous determination of the strength of incentives and the base salary. Their main finding is that incentive pay and reciprocal motivations are substitutes, which is qualitatively in line with our results. Despite these similarities, there are some crucial differences. Most important, Englmaier and Leider assume that positive reciprocity is automatically induced when agents expect to receive a rent. By contrast, in our model reciprocity is not conditional on receiving a rent, but on the belief concerning the principal's type. As a consequence, the principal has to manipulate the agent's belief, implying that, indeed, reciprocal workers earn a rent in equilibrium. Thus, a distinguishing feature of my model is the presence of incomplete information, leading to strategic considerations. These considerations underly my finding that better gift-exchange relations do not always require a higher total expected compensation, as assumed by Englmaier and Leider. Another difference is that Englmaier and Leider restrict their study to contracts that implement an exogenously given (discrete) effort level at the lowest costs, whereas in my model the optimal effort level is endogenously determined by trading off the costs of incentives against the benefits of effort, conditional on the worker's reciprocal tendencies.

Another paper that studies the relation between reciprocal motivations and incentive pay is Dur, Non, and Roelfsema (2008). 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, Non and Roelfsema (2008), the manager's gift consists of management attention. Another simplification is the absence of incomplete information. Both limitations are addressed by Dur (2009), but his model does not allow for incentive pay.

A working paper by Arbak and Kranich (2005) is closest to mine in the sense that incomplete information and Levine-type conditional altruism are key features of their modeling set-up. However, they do not examine how the base salary and incentives interact in credibly signaling the principal's altruism. Moreover, they assume that a limited-liability constraint is always binding, whereas I require that the contract satisfies the agent's participation constraint. Also, while they assume that agents are risk-neutral, in my model agents are risk averse. Finally, by incorporating heterogeneity of worker types I extend the analysis to study the possibilities for screening.

Another related paper is Bellemare and Shearer (2009), who develop a theoretical model of gift-exchange to generalize their experimental results. Interestingly, they study the optimal composition of a gift in terms of an increase in the base salary or the piece rate. They find that an increase in the piece rate is superior to an equivalent increase in 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. The difference in results follows from the difference in assumptions. Bellemare and Shearer do not allow for strategic uncertainty and risk aversion and assume that a limited-liability constraint is always binding

Our paper is also connected to experimental studies on other incentives than rewards, such as imposing a minimum effort requirement or a fine to be

[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.

paid in case of verified shirking. 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". However, as noted above, these results do not extend to incentives that are framed as a reward. From a gift-exchange perspective, distrust is apparently not the only inference workers can make when confronted with explicit incentives. Nevertheless, my paper is in line with some noteworthy results in this literature. I find that incentives themselves do not inhibit gift-exchange, it is the combination of salary and incentives that matters in establishing gift-exchange relations. Interestingly, Falk and Kosfeld (2006) run a treatment that allows the principal to decide not only on installing a minimum effort level, but also on the amount of salary to be paid. In this treatment, higher wages lead to higher average effort levels irrespective of the principal's decision to control. Moreover, Fehr and List (2004)'s study on the relation between fines and trust shows that when a punishment option is available but the principal deliberately refrains from using it, agents exhibit much more trustworthy behavior than when a punishment option is not available. Both results illustrate that we cannot explain the effect of explicit incentives when studied in isolation.

Another topic I address is how contracts can select reciprocal worker types and at the same time signal the principal's altruism. Although I am the first to delve into this issue, my model is not unique in designing a contract that fulfills the dual role of signaling to and screening of the other contracting party. Sliwka (2007) also studies contracts that signal the principal's private information and at the same time screen workers. 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 may also screen worker types when selfish and altruistic workers differ in their 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

3.1 Description of the model.

I consider a risk-averse worker and a risk-neutral principal. The worker is conditionally altruistic, meaning that the extent to which he is altruistic depends on the principal's altruism. The expected utility of a worker of type i is described by:

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

This specification is widely used to describe the utility of risk-averse agents, where r captures the extent of risk-aversion. The production function 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 costs of effort exerted by the worker are represented by $\frac{1}{2}\theta e^2$, and $\gamma_i(\widehat{\alpha}_j)$ represents the extent of worker's altruism towards his principal, which positively depends on the worker's belief about the principal's altruism $\widehat{\alpha}_j$. Workers differ in their altruism function $\gamma_i(\widehat{\alpha}_j)$, where $0 \leq \gamma_i(\widehat{\alpha}_j) < 1$. 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 $\gamma_i(\widehat{\alpha}_j)$. I refer to type k as being more altruistic than type i if and only if $\gamma_k(\widehat{\alpha}_j) \geq \gamma_i(\widehat{\alpha}_j)$ for all $\widehat{\alpha}_j$, where the inequality should be strict for at least some $\widehat{\alpha}_j$. Note that the agent cares about the principal's expected payoff $E(\pi_j)$ instead of his actual payoff. The reason is that it would be nonsensical to assume that the worker is risk-averse over the payoff of a risk-neutral principal. 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 for convenient transformation

[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). He studies how the combination of price and warranty may signal a product's quality and simultaneously screens customers on their willingness to pay for warranty. However, an important difference is that Soberman (2003) abstracts from moral hazard problems: buyer's actions are limited to choosing their preferred price-warranty bundle, and do not otherwise influence the seller's payoff. By contrast, in our model moral hazard is the primary reason for screening, and the contract moreover serves to alleviate the moral hazard problem.

of the utility function into:[5]

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

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.[6] Analogous to the workers, I assume that $0 \leq \alpha_j < 1$. Principals differ in their type α_j, implying that in a separating equilibrium the worker's altruism depends on the specific match between worker and principal.[7] I also refer to the worker's altruism as the quality of the relation.

The timing of the game is as follows. First, the principal decides on a remuneration scheme (b, s). The 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, implying non-employment in equilibrium. Also, I assume that $\overline{u}$ does not depend on $\gamma_i(\widehat{\alpha})$: the worker has no altruistic feelings towards an employer if he does not work for that employer. Finally, the worker decides on his optimal effort level e. I make the standard assumption that effort is non-contractible.

[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 workers welfare compared to the utility suggested by (2). Because we focus on separating equilibria with two types of principals, there is no uncertainty on the principal's type in equilibrium, and hence it is safe to ignore the worker's 'preference for certainty' for simplicity. Taking the effect 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] Obviously, the coefficients α_j and $\gamma_i(\widehat{\alpha}_j)$ 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 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. Confining altruism to both player's material payoff does not affect the qualitative results, but is inconvenient analytically.

[7] When types are observable, it would be possible to assume that the principal's altruism depends on his beliefs about the worker's altruism, i.e. α_j as a function of $\gamma_i(\alpha_j)$. However, when types are unobservable, such an assumption requires that higher-order beliefs are formed. When the principal would be conditionally altruistic, a worker's degree of altruism would not only depend on his own type and on his beliefs concerning the principal's type, but also on what he believes to be the principal's beliefs. The latter determine the principal's actual altruism after all.

3.2 Analysis when types are observable

In this section, I assume that both players learn about each other's type before they make any decision. I solve for a subgame perfect equilibrium using backward induction. The worker's effort choice follows from maximization of his utility function (2), which yields the following first-order condition:

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

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 ($b = 1$), his actions do not affect the principal's profits. Therefore, his choice of effort is independent of his altruism: any worker type equates the marginal costs of effort with the marginal product ($\theta e = 1$). Rewriting the first-order condition (4) gives the worker's optimal 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: 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.[8] 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.[9]

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

[8] Recall that 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).

[9] A second effect would arise if I would allow for a cost of effort function that has a positive third-order derivative. The more altruistic the worker, the more effort he already exerts out of an altruism motive, which reduces his responsiveness to incentives because additional effort becomes increasingly costly. Similarly, incentives reduce the responsiveness to gift-exchange, which explains the crowding out in Bellemare and Shearer (2009)'s model.

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. Hence, we can insert the base salary implied by the participation constraint into the objective function:

$$\max_{b} \;\; E(\pi_j) = (1-b)e_i^* - \left[\overline{u} - be_i^* + \frac{1}{2}\theta e_i^{*2} + \frac{1}{2}r\sigma^2 b^2 - \gamma_i(\alpha_j)E(\pi_j)\right] + \alpha_j \overline{u}.$$

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 benefits the principal with the size of the marginal product. However, because the worker's participation constraint is binding, the worker needs to be compensated for the marginal cost of providing effort. Further, risk-averse workers need to be compensated for exposure to income uncertainty. Clearly, when workers are risk-neutral ($r = 0$), the first-order condition implies that the marginal benefits of effort are equated with the marginal costs. We derive the payoff-maximizing bonus b^* by inserting (5) and its derivative into the first-order condition:

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

Clearly, when workers are risk averse, the bonus decreases in the quality of the relation $\gamma_i(\alpha_j)$. 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 giving a 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 employer

$(\gamma_i(\alpha_j) = 0)$, we obtain the standard solution: the bonus is decreasing in the amount of 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. On the one hand, additional worker motivation has a positive effect on effort, but this is possibly more than offset by the corresponding decrease in financial incentives.[10] 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 ($b^* > \frac{1}{2}$), and therefore the reduction in incentives is quantitavely unimportant compared to the increase in altruistic feelings.

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

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

As long as the principal's payoff is positive[11], total compensation is decreasing in the quality of the relation. The main reason is a compensating wage differential: the more utility workers derive from the non-monetary aspects of their jobs, the lower the required monetary compensation. A second reason is that the strength of financial incentives decreases in the quality of the relation, which is reflected in a lower risk-premium when workers are risk averse. Thus, the analysis leads to the following result:

Proposition 1 *When types are observable, good relations are associated with weaker incentives and lower total monetary compensation.*

4 Unobservable types: Signaling and screening

In the previous section I assumed that the quality of the relation is determined before the contract is written. Although this is a reasonable assump-

[10] The total derivative $\frac{de_i^*}{d\gamma_i(\alpha_j)}$ is described by

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

which is positive if $\theta r\sigma^2 > [1-\gamma_i(\alpha_j)]^2$.

[11] Firms can only stay in business as long as profits are at least zero. It is easily verified that profits are increasing in relational capital, and so every principal who is more altruistic than the least altruistic principal in the population obtains a positive payoff.

tion when workers have been working at the firm for some time or when reputational concerns are important, in this section I relax this assumption and assume that neither the worker's type nor the principal's type is observable. Because types are private information, contracts potentially have a dual role of both signaling the employer's type and screening worker's types. I confine this study to separating equilibria, because in these equilibria different contracts are offered by different types, which leads to predictions on the how compensation policies relate to the quality of manager-worker relations. By contrast, in pooling equilibria, contracts are independent of the principal's and agent's type. I now first introduce some simplifying assumptions. Then, I study a separating equilibrium where the principal signals his altruism, but abstains from screening worker types. Finally, I consider a separating equilibrium where the altruistic principal writes a contract that signals his altruism and simultaneously screens worker types.

4.1 Simplifying assumptions

The first step in simplifying the problem is to restrict the number of types. I assume that there are only two types of workers and employers, namely selfish (l) and (conditionally) altruistic (h) types. A selfish principal has no altruistic feelings at all: $\alpha_l = 0$. Because the principal's type is unobservable, the agent's altruism depends on his beliefs concerning the principal's type, denoted $\widehat{\alpha}_j$. When applicable, I rule out unreasonable beliefs by requiring that beliefs satisfy the intuitive criterion.[12] In addition, I make some simplifying assumptions on the altruism functions of the worker. Specifically, I assume that a selfish worker type never takes the principal's welfare into account regardless of the principal's altruism ($\gamma_l(\widehat{\alpha}_j) = 0$ for any $\widehat{\alpha}_j$) and that a reciprocal worker type is completely egoistic when he believes that the principal is selfish ($\gamma_h(\widehat{\alpha}_l) = 0$). To make the problem non-trivial, I require that $\gamma_h(\widehat{\alpha}_h) > 0$. I assume that a fraction δ of all workers is selfish, whereas the remaining fraction $(1 - \delta)$ is conditionally altruistic. It is not necessary to introduce notation for the fraction of altruistic principals.[13] Finally, I assume that there are more reciprocal workers than altruistic principals, which

[12] 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).

[13] This does not mean, however, that the fraction of altruistic principals is irrelevant. Specifically, it determines whether it is worthwhile to coordinate on a separating equilibrium. The larger the fraction of altruistic principals, the more attractive to coordinate on a pooling equilibrium instead. Moreover, when there are more altruistic principals than reciprocal workers, a signaling and screening equilibrium does not exist.

prevents competition for reciprocal workers when altruistic principals engage in screening.

4.2 Analysis: Signaling altruism

In this section I study a separating equilibrium where the altruistic principal writes a contract that signals his altruism. 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 payoff (see equation (3)) can be written as:

$$E(\pi_l) = (1 - b_l) \left[\delta e_{ll} + (1 - \delta) e_{hl}\right] - s_l, \tag{7}$$

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.[14] 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:

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

Clearly, because $\gamma_i(\widehat{\alpha}_l) = 0$ irrespective of the worker's type, both worker types exert the same effort and derive the same utility from accepting the selfish type's equilibrium contract. This implies that both types require the same compensation to satisfy their participation constraint. Because in any fully separating equilibrium a selfish principal obviously has no reason to signal his type, he does not distort his optimal contract choice compared to the case when types are observable. Thus, the selfish principal will offer a bonus $b_l = \frac{1}{1+r\sigma^2\theta}$ and a base salary that exactly satisfies the worker's participation constraint:

$$s_l = \overline{u} - b_l e_{il} + \frac{1}{2}\theta e_{il}^2 + \frac{1}{2} r\sigma^2 b_l^2. \tag{8}$$

In order to derive the optimal contract choice of the altruistic principal, it is instructive to inspect his payoff function. Assuming that the contract

[14] Because in equilibrium beliefs are always based on the observed contract offer (b_l, s_l) or (b_h, s_h), this shorthand notation suffices to describe the equilibrium contracts.

(b_h, s_h) succeeds in credibly signaling the principal's altruism, equation (3) can be rewritten to:

$$E(\pi_h) = (1 - b_h) [\delta e_{lh} + (1 - \delta) e_{hh}] - s_h + \alpha_h [\delta E(u_l) + (1 - \delta) E(u_h)] . \quad (9)$$

This equation shows that the altruist's payoff positively depends on the fraction of reciprocal worker types 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 valuable for a principal who has altruistic feelings. For these two reasons, the altruistic principal may benefit from writing a contract that convinces the reciprocal worker types that he is an altruist. To accomplish this, the equilibrium contract (b_h, s_h) should satisfy two incentive compatibility constraints (ICC): the selfish principal should have no incentive to mimic the altruist and vice versa:

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

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

where $e_{\delta l} = \delta e_{ll} + (1 - \delta) e_{hl}$ and $e_{\delta h} = \delta e_{lh} + (1 - \delta) e_{hh}$. It is essential to note that when ICC1 is satisfied, ICC2 can only be satisfied if the difference $\alpha_h [\delta E(u_l) + (1 - \delta)E(u_h)] - \alpha_h \overline{u}$ is large enough. This observation reveals why in equilibrium a principal with altruistic feelings is willing to engage in costly signaling: because he enjoys worker's utility to some extent, his payoff increases when a reciprocal worker believes him to be an altruist. Thus, an altruistic principal is willing to signal his type because he values good relations in the workplace. By contrast, the aforementioned observation reveals that the resulting difference in effort $(e_{\delta h} - e_{\delta l})$ is no motive for signaling in equilibrium. The reason is that a selfish principal will imitate any contract that yields a higher monetary payoff than his own equilibrium contract. Therefore, the difference in effort $(e_{\delta h} - e_{\delta l})$ must be reflected in higher payments to the worker.

Reasoning further along these lines, it is straightforward to show that ICC1 is always binding. When the altruistic principal ignores ICC1, he chooses the same contract as when types are observable. Such a contract clearly satisfies both workers' participation constraint and ICC2. However, such a contract violates ICC1: because profits are increasing in the worker's altruism, the egoistic principal finds it highly profitable to mimick the altruistic employer. Therefore, ICC1 must be binding in equilibrium, implying that the optimal contract will be a (b_h, s_h) −combination such that ICC1 holds with equality and that ICC2 is slack.[15]

In addition to the two incentive compatibility constraints, the altruistic principal's equilibrium contract should satisfy both workers' participation contraint (PC).[16] 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, the reciprocal worker's PC is described by

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

Comparison of these two constraints reveals that reciprocal workers derive more utility from a given equilibrium contract than selfish workers. This implies that when the selfish worker's PC is satisfied, the 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 altruistic principal's objective is to write a contract that maximizes his expected payoff, provided the four constraints outlined above are satisfied. Because PCH and ICC2 are both slack, I can reformulate the problem in a convenient way. Since 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 his worker. The maximization problem can thus be formulated as:

$$\max_{s_h, b_h} \quad \delta E(u_l) + (1-\delta) E(u_h)$$

$$\text{s.t.} \quad s_h = (1-b_h) e_{\delta h} - E(\pi_l), \tag{ICC1}$$

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

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

[15] 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. This condition is always satisfied except when $r\sigma^2 = 0$. However, for tractability we assume that $r\sigma^2 > 0$.

[16] In this section, we assume that both worker types must be willing to work for an altruistic employer or a selfish employer, i.e. $E(u_i) \geq \overline{u}$ for both types. As we will see in the next section, the altruistic employer may find it optimal to offer a contract that selects the reciprocal workers only.

assume that PCL is satisfied and confirm afterwards that our conjecture is correct. This allows us to rewrite the problem by substitution of $E(u_l)$, $E(u_h)$, and s_h, yielding the following:

$$\begin{aligned}\max_{b_h} \quad & \delta\left(e_{lh} - \frac{1}{2}\theta e_{lh}^2 - \frac{1}{2}r\sigma^2 b_h^2\right) + \\ (1-\delta) & \left(e_{hh} - \frac{1}{2}\theta e_{hh}^2 - \frac{1}{2}r\sigma^2 b_h^2 + \gamma_h(\widehat{\alpha}_h)E(\pi_h)\right) - E(\pi_l).\end{aligned}$$

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. Therefore, depending on the fraction of selfish workers in the population (δ), the payoff-maximizing bonus b_h lies between $b^* = \frac{[1-\gamma_h(\widehat{\alpha}_h)]^2}{[1-\gamma_h(\widehat{\alpha}_h)]^2+\theta r\sigma^2}$ 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:

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

Thus, the best an altruistic principal can do is to offer the bonus that maximizes the joint surplus and increase the base salary up to the point that the selfish principal is not any longer willing to imitate.

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 larger 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 total expected 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 necessitate paying higher compensation. As shown in the observable types case, a good relation does not automatically lead to more effort, because the intensity of financial incentives decreases in the quality of the relation, i.e. $b_h < b_l$. Therefore, we cannot be sure that altruistic workers exert more effort when employed by an altruistic principal, whereas selfish workers unambiguously provide less effort. Thus, paying a sufficiently high total expected compensation will always be interpreted as a signal of altruism, but a low total compensation does not necessarily disprove altruism as long as it is accompanied by weak financial incentives.

Proposition 2 *In a signaling equilibrium, good relations are associated with weaker incentives, but not necessarily with higher total compensation.*

The equilibrium is illustrated by figure 1.[17] The equations underlying the picture are presented in appendix B. The figure shows ICC1 for two different values of $\gamma_h(\widehat{\alpha}_h)$ and the participation constraints PCL and PCH. Thus, ICC1 represents an isoprofit curve that indicates the minimum base salary required to keep the selfish principal from imitating.[18] Similarly, PCL and PCH are indifference curves representing the lowest base salary that is acceptable to selfish and reciprocal workers respectively. The arrows thus indicate the area of feasible contracts. The dotted line represents a weighted average of the indifference curves, thus representing a hypothetical average worker. The optimum is where the indifference curve of the hypothetical average worker is tangent to ICC1. The corresponding bonus maximizes the joint surplus.

For the remainder of the paper, it is important to understand the intuition behind the curves. The slope of ICC1 depends on $\gamma_h(\widehat{\alpha}_h)$: it has an inverted u-shape provided $(1-\delta)\gamma_h(\widehat{\alpha}_h) < \frac{1}{2}$, but is always decreasing in the bonus when $(1-\delta)\gamma_h(\widehat{\alpha}_h) > \frac{1}{2}$. 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 the effort of highly reciprocal workers is relatively high and insensitive to incentive pay, the latter effect dominates when $(1-\delta)\gamma_h(\widehat{\alpha}_h) > \frac{1}{2}$. Both PCL and PCH may slope downwards or upwards, the latter case being depicted in figures 1 and 2. The reason is that an increase in the expected bonus payment has an ambiguous effect on worker's utility. On the one hand, the additional payment benefits the worker, implying that the base salary should decrease

[17]The equilibrium exists for all parameter values, provided $\gamma_h(\widehat{\alpha}_h) > 0$ and $r\sigma^2 > 0$. All other equilibria can be ruled out by applying the intuitive criterion.

[18]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.

to keep expected utility constant. On the other hand, an increase in the bonus exposes the worker to more risk, for which he must be compensated. The latter effect dominates when $r\sigma^2\theta > 1$. 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 contract than an egoist. Moreover, PCH is more inclined to slope downwards than PCL. The reason is that the altruistic worker always exerts more effort than the egoist, implying that a given increase in the bonus leads to a larger increase in expected payments for the altruistic worker than for the egoistic worker. Therefore, a larger decrease (or a smaller increase) in the base salary is required to keep 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 employer, 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 employer 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), including the selfish principal's offer (b_l, s_l). Since for a given bonus 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. There are three considerations why the selfish worker's utility may be higher or lower than his reservation utility. First, the selfish worker benefits from the reciprocal worker's productivity: the latter's relatively high productivity must be reflected in the base salary to ensure that altruistic principal's expected monetary pay-off does not exceed the selfish principal's profits. This free-rider effect is manifest in figure 1 by the fact that for $b_h = b_l$, $s_h > s_l$. Second, the bonus is not set optimally according to the preferences of the selfish worker type: the bonus b_h compromises the preferences of both worker types, as shown above. Third, the bonus affects the distribution of the total surplus between the selfish and reciprocal worker. Because a selfish worker provides less effort than a reciprocal worker, changes in the bonus have a smaller effect on the selfish worker's utility. As a consequence, the fact that $b_h < b_l$ benefits the selfish worker relative to the reciprocal worker. This can be seen in figure 1 from the fact that the PCL is steeper than PCH, implying that the selfish type gains from a decrease in the bonus and a corresponding change in the base salary that keeps the altruist's utility unchanged. I prove in appendix A that the third effect dominates the second effect, and hence that PCL is always satisfied.

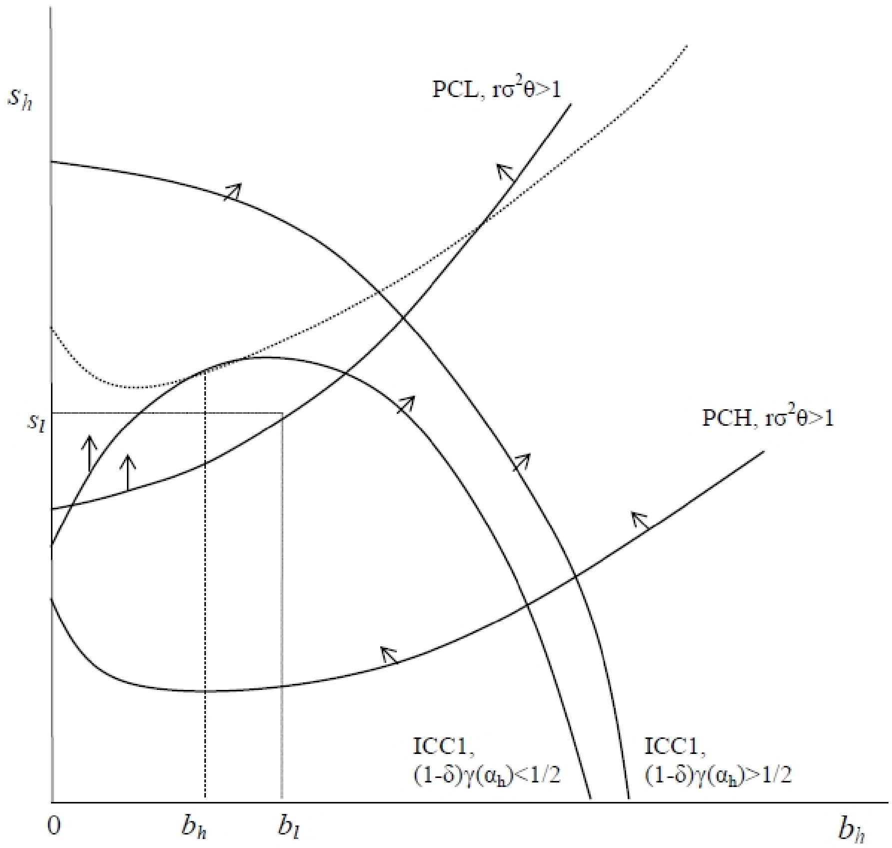

Figure 1

4.3 Analysis: Signaling altruism and screening workers

In this section I study a separating equilibrium where the altruistic principal writes a contract that signals his altruism and simultaneously screens worker types. The analysis proceeds along the same lines as in the previous section. I start the analysis by deriving the egoistic principal's contract choice. Because all workers are completely egoistic when they believe that the principal is selfish (i.e. $\gamma_i(\widehat{\alpha}_l) = 0$), it is easily verified that a selfish principal has no signaling or screening motive in equilibrium. Lacking signaling and screening motives, a selfish principal does not distort his contract choice compared to the case when types are observable. Thus, I easily obtain the same result as in the previous section: the selfish principal will offer a bonus $b_l = \frac{1}{1+r\sigma^2\theta}$ and the smallest base salary workers are willing to accept.

Deriving the altruistic principal's contract choice is more complicated. As shown in the previous section, the altruistic principal's utility is increasing in the fraction of reciprocal workers, provided he convinces them that he is an altruist. Therefore, he may benefit from writing a contract that selects reciprocal workers only. Screening of worker types can be accomplished by offering a contract (b_h, s_h) that is unacceptable to selfish worker types, but is attractive for reciprocal workers. Since both workers expect to obtain their reservation utility $\overline{u}$ when they accept the selfish principal's contract, screening requires that the contract simultaneously violates PCL and satisfies PCH. For ease of exposition, I refer to violating PCL as satisfying the screening constraint (SCC):

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

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:

$$(1 - b_l)e_{ll} - s_l \geq (1 - b_h)e_{hh} - s_h \tag{ICC1'}$$

$$(1 - b_l)e_{ll} - s_l + \alpha_h \overline{u} \leq (1 - b_h)e_{hh} - s_h + \alpha_h E(u_h) \tag{ICC2'}$$

Since the structure of the problem is unchanged compared to the previous section, 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. Moreover, we observed in the previous section that when the principal abstains from screening workers, PCL is always satisfied. Therefore, succesful screening requires that SCC is explicitly taken into account and hence is also binding. As in the previous section, a reasonable conjecture is that PCH is irrelevant: because the altruistic principal's monetary payoff is constrained by ICC1', the best he can do is to maximize the reciprocal's worker's utility (taking SCC into account). Again I will check afterwards whether our conjecture is correct. I can write the problem as follows:

$$\max_{b_h, s_h} \quad E(u_h) = b_h e_{hh} + s_h - \frac{1}{2}\theta e_{hh}^2 - \frac{1}{2} r\sigma^2 b_h^2 + \gamma_h(\widehat{\alpha}_h) E(\pi_h)$$

$$\text{s.t. } s_h = (1 - b_h)e_{hh} - E(\pi_l) \tag{ICC1'}$$

$$\text{s.t. } b_h e_{lh} + s_h - \frac{1}{2}\theta e_{lh}^2 - \frac{1}{2} r\sigma^2 b_h^2 \leq \overline{u} \tag{SCC}$$

Since both ICC1' and SCC are binding, I conclude that the equilibrium is at an intersection point of ICC1' and SCC. This also proves that the reciprocal worker's PC is satisfied, because when SCC is binding, PCH is slack.

Figure 2 illustrates that the equilibrium is at an intersection of ICC1' and SCC. 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 altruistic worker's indifference curve (PCH) up, it can easily be seen that, given the constraints, his expected utility is maximized at an intersection of ICC1' and SCC, specifically the intersection that specifies $b_h > b_l$.

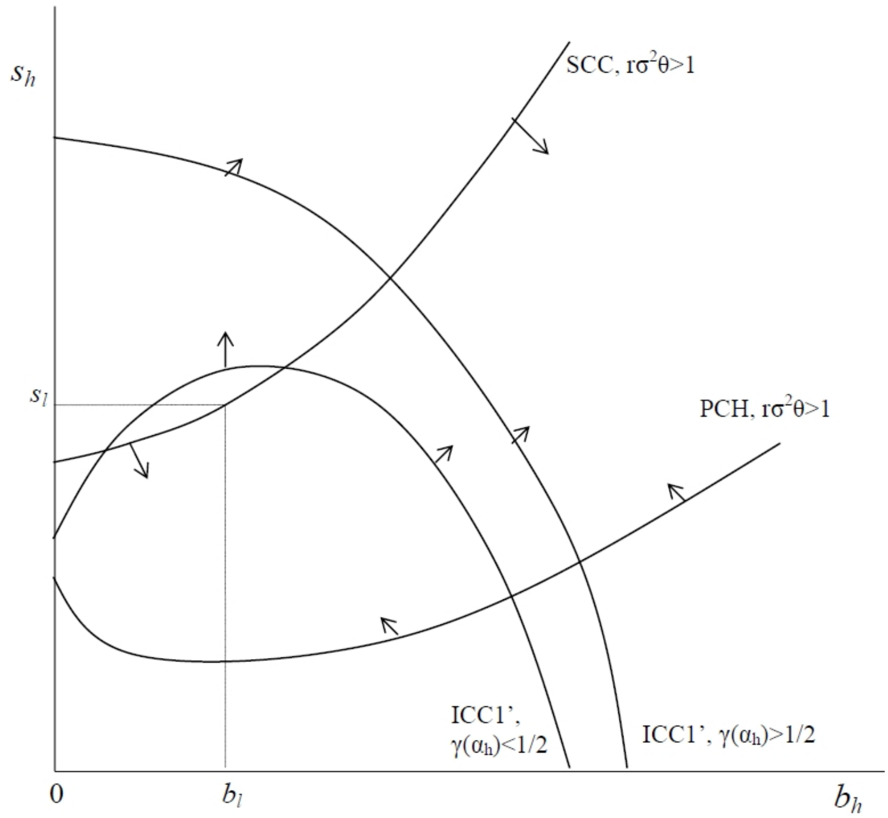

Figure 2

It can be shown that this is a more general result. In what follows, I will show that an equilibrium always exists and that it is characterized by a bonus b_h on the interval $(b_l, 1)$. To prove existence, first recall that for $b_h = b_l$, ICC1' lies above SCC. It can be shown that 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 agent is the full residual claimant and SCC thus specifies the 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.[19] Similarly, because the altruistic principal is no residual claimant, ICC1' specifies the maximum possible amount the principal receives 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.[20] The amount the worker is willing to pay for the firm (SCC) is always smaller than the 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. Noting that when (ICC1') and (SCC) are expressed as a base salary, both amounts are negative, we thus conclude that an intersection point on the interval $(b_l, 1)$ always exists.

When there are two intersection points, the principal prefers the bonus at the intersection point on the interval $(b_l, 1)$. Consider the case when $\gamma_h(\widehat{\alpha}_h)$ is small, implying that ICC1' has an inverted u-shape (see figure 2). The principal prefers the intersection point that specifies the highest bonus, because an increase in the bonus is more beneficial for the reciprocal worker than the egoistic worker due to the former's higher effort. The reciprocal worker's higher effort is reflected in his indifference curve (PCH) that is more inclined to slope downwards. Moving along an egoistic worker's indifference curve (SCC) therefore increases the reciprocal worker's expected payoff. Thus, the principal prefers the equilibrium contract at the intersection on the interval $(b_l, 1)$. When the SCC is downward sloping, a second intersection point may also exist when $\gamma_h(\widehat{\alpha}_h)$ is sufficiently high and $b_h > 1$. However, I show in appendix A that the intersection point on the interval $(b_l, 1)$ is strictly preferred. Finally, the total expected compensation earned/paid by the al-

[19] To see this, let s_h^{SCC} denote the maximum salary the SCC allows for. The selfish worker's willingness to pay $(-s_h^{SCC})$ can be found by rewriting the SCC and inserting $b_h = 1$:

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

[20] Let $s_h^{ICC1'}$ denote the base salary that keeps the egoistic principal from imitating. 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 selfish agent's profits. Using equations (7) and (8), 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, $E(\pi_l) > -s_h^{SCC}$, implying that $s_h^{ICC1'} < s_h^{SCC}$.

truistic worker/principal is always larger than that of the selfish type.[21] The reason is that because the reciprocal worker unambiguously provides more effort in equilibrium ($e_{hh} > e_{ll}$), the altruistic principal has to pay more to ensure that his expected monetary payoff does not exceed the selfish principal's profits. To summarize:

Proposition 3 *In a signaling and screening equilibrium, good relations are associated with stronger incentives and larger total compensation.*

These results stand in remarkable contrast with the preceding results. The reason for these diverging findings is that screening can (most 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'. Because the reciprocal worker faces stronger incentives than the selfish worker, he unambiguously provides more effort, implying that the altruistic principal has to pay more than the selfish principal to discourage him from imitating.

One may wonder whether the altruistic principal prefers this signaling and screening equilibrium above the signaling equilibrium. Because his profits are identical in the two situations (namely the same as the selfish principal's payoff), the equilibrium that yields the largest immaterial payoff (or, equivalently, the highest average worker utility) is preferred. Screening of worker types has the advantage that only altruistic worker types are attracted, which has a positive effect on average worker utility. However, screening is also costly: the bonus is distorted compared to the efficient bonus level. Clearly, incurring the costs of screening is unattractive when the large majority of workers is altruistic, but attractive when the fraction of altruistic workers is small. I provide a formal proof in appendix A.

5 Concluding remarks

I have studied the relation between gift-exchange, the power of incentives, and worker heterogeneity in an otherwise standard principal-agent model. Following Levine (1998), I have assumed that some agents care more for their principal when they are convinced that the principal cares for them.

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

$$b_h e_{hh} + s_h = b_l e_{ll} + s_l + e_{hh} - e_{ll},$$

implying that $b_h e_{hh} + s_h > b_l e_{ll} + s_l$ since $e_{hh} > e_{ll}$.

Abstracting from the incomplete information problem, I have found that good manager-worker relations substitute for incentives and allow for a reduction in total expected compensation, as the quality of the relation is reflected in a compensating wage differential. Assuming that types are private information, I have found that principals signal their altruism by offering relatively weak incentives, but not necessarily by paying a relatively high total expected compensation. The key to understanding this result is that the bonus and the fixed salary simultaneously provide cues about the principal's altruism. Expecting that a gift-exchange relation will be established, altruistic principals provide relatively weak incentives because the agent's and principal's preferences are more aligned under gift-exchange. As a result of the low incentive-intensity, gift-exchange may be associated with relatively low productivity, and therefore a relatively low compensation may already suffice to distinguish the altruistic principal from the selfish type. Finally, assuming that some agents 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 worker's altruism, because altruistic workers put in more effort than selfish workers and hence gain more from output-contingent pay. Thus, strong incentives are offered to attract the workers who need them least.

Although the model is highly stylized, the results can be generalized to less restrictive settings. First, the relevance of the observable types case may well extend beyond an instructive benchmark case. Although full observability is a strong assumption, workers typically receive highly informative signals about the employer's care for them regardless of the information that is contained in the wage contract. Thus, I am confident that the gist of the results extends to a setting where workers receive a sufficiently informative signal of the principal's type. Second, for ease of exposition I assumed that all workers are completely egoistic when they believe that the principal is selfish, and that a selfish worker never takes a principal's payoff into account. Relaxing these assumptions complicates the analysis, but does not affect the results qualitatively. The main complication is that selfish principals also have an incentive to screen when some workers care about the principal's welfare, even if they believe that the principal is selfish. Therefore, selfish principals will offer two different contracts, yielding the same level of profits. Third, the qualitative results do not change when there is free entry of selfish employers. Finally, relaxing the assumption that all workers have the same reservation utility does not affect the qualitative results when types are private information. When types are observable, reciprocal worker's better outside option may (but need not) induce the principal to pay them higher wages.

An empirical implication of the model is that the association between total expected compensation and the quality of manager-worker relations depends on the worker's uncertainty about the principal's type. To the extent that worker's uncertainty can be proxied by tenure, a positive relation between total compensation and the quality of the relation is more likely for less tenured workers. The model also predicts that gift-exchange relations are associated with weaker incentives. Unfortunately, most employment surveys are silent on the strength of incentives an individual receives. They only provide information on *whether* workers receive bonus pay. It is straightforward, however, to reinterpret the model by assuming that incentive pay can only be implemented at a fixed cost.[22] The empirical prediction then becomes that firms with good relations are less likely to implement a pay-for-performance scheme. The reason is that the implementation of a costly incentive scheme is less worthwhile when managers can also rely on altruism as a motivational force.

Of course, there are several 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. Previous studies suggest that effects of these policies on gift-exchange relations cannot be studied in isolation. Hence, it may be worthwhile to investigate these issues further. A second point is that our Levine-type approach of modeling altruism ignores fairness considerations in the sense that people may have some idea of what constitutes a fair income, and reciprocate when income falls below. Thus, it may not be sufficient that the contract distinguishes an altruistic principal from a selfish principal, the compensation should also be perceived as fair. This could preclude a compensating wage differential, although the notion of fair income may take the quality of the relation into account. Third, the existence of a signaling and screening equilibrium crucially depends on the assumption that the number of altruistic principals does not exceed the number of reciprocal workers. By contrast, when altruistic principals compete for reciprocal workers, the signaling equilibrium always gives a higher payoff and hence is the only equilibrium that survives the intuitive criterion. Moreover, a signaling equilibrium does not exist when an altruistic principal can increase his

[22] It should be noted that the signaling and screening equilibrium cannot be reinterpreted so easily. When introducing a bonus is not worthwhile for the selfish principal, the gains from screening must be extremely large to induce the selfish principal to introduce a bonus.

payoff by deviating to a screening contract. Finally, 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 such as wage bargaining with trade unions impact on the prospects for establishing gift-exchange relations. Thus, there is ample room for further research, both theoretically and empirically.

6 Appendix A: formal proofs

6.1 Proof that PCL is always satisfied.

This appendix contains the proof that in a signaling equilibrium, the selfish worker's PC is always satisfied. First note that when $\gamma_h(\widehat{\alpha}_h) = 0$ both principals offer the same contract and hence both worker types obtain their reservation utility $\overline{u}$. By taking the derivative to $\gamma_h(\widehat{\alpha}_h)$, I show that the selfish worker's expected utility $E(u_l)$ increases in $\gamma_h(\widehat{\alpha}_h)$ and hence always exceeds his reservation utility.

$$\frac{dE(u_l)}{d\gamma_h(\widehat{\alpha}_h)} = \frac{db_h}{d\gamma_h(\widehat{\alpha}_h)}\left[\frac{de_{lh}}{db_h}(b_h - \theta e_{lh}) + e_{lh} - r\sigma^2 b_h\right] + \frac{ds_h}{d\gamma_h(\widehat{\alpha}_h)}. \tag{A1}$$

Because ICC1 is binding, $\frac{ds_h}{d\gamma_h(\widehat{\alpha}_h)}$ is defined as:

$$\begin{aligned}\frac{ds_h}{d\gamma_h(\widehat{\alpha}_h)} &= (1-b_h)(1-\delta)\frac{de_{hh}}{d\gamma_h(\widehat{\alpha}_h)} + \\ \frac{db_h}{d\gamma_h(\widehat{\alpha}_h)}&\left[(1-b_h)\left(\delta\frac{de_{lh}}{db_h} + (1-\delta)\frac{de_{hh}}{db_h}\right) - e_{\delta h}\right]\end{aligned} \tag{A2}$$

Combining these expressions and applying the envelop theorem (utility maximization implies that $b_h - \theta e_{lh} = 0$), we obtain:

$$\begin{aligned}\frac{dE(u_l)}{d\gamma_h(\widehat{\alpha}_h)} &= (1-\delta)(1-b_h)\frac{de_{hh}}{d\gamma_h(\widehat{\alpha}_h)} + \frac{db_h}{d\gamma_h(\widehat{\alpha}_h)}\left[(1-\delta)(e_{lh} - e_{hh}) - r\sigma^2 b_h\right] \\ &+(1-b_h)\frac{db_h}{d\gamma_h(\widehat{\alpha}_h)}\left[\delta\frac{de_{lh}}{db_h} + (1-\delta)\frac{de_{hh}}{db_h}\right]\end{aligned}$$

This can be rewritten further to elicit the three effects described in the main text:

$$\begin{aligned}\frac{dE(u_l)}{d\gamma_h(\widehat{\alpha}_h)} &= (1-\delta)(1-b_h)\frac{de_{hh}}{d\gamma_h(\widehat{\alpha}_h)} + \delta\frac{db_h}{d\gamma_h(\widehat{\alpha}_h)}\left[\frac{de_{lh}}{db_h}(1-b_h) - r\sigma^2 b_h\right] \\ &+ (1-\delta)\frac{db_h}{d\gamma_h(\widehat{\alpha}_h)}\left[\frac{de_{hh}}{db_h}(1-b_h) - r\sigma^2 b_h + (e_{lh} - e_{hh})\right].\end{aligned}$$

The first term is a pure 'free-rider effect': because ICC1 is binding, the selfish worker benefits to some extent from the additional effort of the reciprocal worker. The second term is negative (since $\frac{db_h}{d\gamma_h(\widehat{\alpha}_h)} < 0$ and the term between brackets is positive) and captures that b_h is not set according to the selfish agent's preferences. The third term is positive (since $\frac{db_h}{d\gamma_h(\widehat{\alpha}_h)} < 0$ and the term between brackets is negative) and reflects the fact that the selfish worker gets a relatively larger share of the surplus when the bonus is reduced. After considerable rewriting (using the first-order conditions for optimal effort and expressions for e_{lh} and e_{hh}), we obtain:

$$\frac{dE(u_l)}{d\gamma_h(\widehat{\alpha}_h)} = (1-\delta)(1-b_h)\left[\frac{db_h}{d\gamma_h(\widehat{\alpha}_h)}\left(-\gamma_h(\widehat{\alpha}_h)\frac{\gamma_h(\widehat{\alpha}_h)}{\theta}\right) + \frac{de_{hh}}{d\gamma_h(\widehat{\alpha}_h)}\right] > 0,$$

which shows that the third effect dominates the second effect when $\gamma_h(\widehat{\alpha}_h) > 0$, and that the total effect on the selfish worker's utility is always positive. Hence, the selfish worker's utility is always higher than his reservation utility. Note that this implies that in a signaling and screening equilibrium, SCC is always binding.

6.2 Proof that an equilibrium where $b_h > 1$ is inferior to $b_h \in (b_l, 1)$.

In this appendix, I show that an intersection point of SCC and ICC1' where $b_h > 1$ gives a strictly lower payoff than when $b_h \in (b_l, 1)$, and can therefore be ruled out by the intuitive criterion. I refer to appendix B for the mathematical expressions used here. Necessary conditions for existence of such an intersection point are that SCC slopes downwards, i.e. $\theta r\sigma^2 < 1$, and that $\gamma_h(\widehat{\alpha}_h)$ is so high that SCC is steeper than ICC1' for $b_h > 1$, i.e. $\frac{ds_h^{ICC1}}{db_h} > \frac{ds_h^{SCC}}{db_h}$. To prove that the altruistic worker and principal prefer the equilibrium where $b_h \in (b_l, 1)$, it suffices to show that PCH has a flatter slope than ICC1' when $b_h > b_l$, implying that the altruistic worker gains when the bonus is reduced along ICC1'. Stated otherwise, I show that $\frac{ds_h^{ICC1}}{db_h} < \frac{ds_h^{PCH}}{db_h}$ for $b_h > b_l$. Inserting the expressions provided in appendix B, I obtain after considerable rewriting:

$$1 - 2\gamma_h(\widehat{\alpha}_h)\left(1 - \frac{1}{2}\gamma_h(\widehat{\alpha}_h)\right)(1-b_h) < b_h\left(1+\theta r\sigma^2\right).$$

It is straightforward to verify that this inequality holds for all $b_h > b_l$. This proves that the altruistic worker prefers the equilibrium on the interval $(b_l, 1)$.

6.3 Conditions under which signaling and screening is attractive.

This appendix shows under what conditions the principal prefers the signaling and screening equilibrium above the signaling equilibrium. First note that when $\gamma_h(\widehat{\alpha}_h) = 0$, the payoff is equal in both equilibria. I now consider how the principal's payoff changes when $\gamma_h(\widehat{\alpha}_h)$ increases in both equilibria, where the change in the principal's payoff is equal to the change in total utility.

In a signaling equilibrium, the effect of a change in $\gamma_h(\widehat{\alpha}_h)$ on total utility is given by $\delta \frac{dE(u_l)}{d\gamma_h(\widehat{\alpha}_h)} + (1-\delta) \frac{dE(u_h)}{d\gamma_h(\widehat{\alpha}_h)}$, or after some rewriting:

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

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

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

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

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 all workers are reciprocal instead of a fraction $(1-\delta)$. The first part is zero if $\gamma_h(\widehat{\alpha}_h) = 0$, but becomes smaller (increases in absolute value) when $\gamma_h(\widehat{\alpha}_h)$ becomes larger. The second part is concave in $\gamma_h(\widehat{\alpha}_h)$, because b_h is increasing and hence $\frac{de_{hh}}{d\gamma_h(\widehat{\alpha}_h)}$ is decreasing in $\gamma_h(\widehat{\alpha}_h)$.

The signaling and screening equilibrium is always preferred for $\gamma_h(\widehat{\alpha}_h)$ sufficiently close to zero. The reason is that reciprocal workers only are attracted. When $\gamma_h(\widehat{\alpha}_h)$ increases, the cost of distorting the bonus becomes more severe, and at some point the signaling equilibrium will be preferred. The larger the fraction of selfish workers δ, the more attractive to screen workers.

7 Appendix B: The constraints and their properties.

This section provides the mathematical expressions for ICC1, ICC1', SCC (or, equivalently, PCL) and PCH that underly the figures, and proves some of their properties described in the main text. Rewriting ICC1 yields:

$$s_h \geq (1-b_h)e_{\delta h} - \left[(1-b_l)e_{\delta l} - s_l\right].$$

I now show that the base salary that keeps the egoist from imitating, denoted s_h^{ICC1}, decreases in the bonus as long as $(1-b_h)e_{\delta h}$ decreases in the bonus. This is always the case when $(1-\delta)\,\gamma_h(\widehat{\alpha}_h) > \frac{1}{2}$, implying that $e_{\delta h}$ is rather large and insensitive to increases in the bonus. Using equation (5) for the worker's effort choice and equation (6) for b_l, I obtain:

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

Inspection of the derivative to b_h proves that the minimum base salary required by ICC1 initially increases in the bonus provided $(1-\delta)\,\gamma_h(\widehat{\alpha}_h) < \frac{1}{2}$, but always decreases in the bonus when $(1-\delta)\,\gamma_h(\widehat{\alpha}_h) > \frac{1}{2}$:

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

The derivation of ICC1' and its derivative to b_h merely requires inserting $\delta = 0$ into the expressions above.

Assuming that the screening constraint (or, equivalently, PCL) holds with equality, SCC can be rewritten to:

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

Clearly, the maximum salary the SCC allows for, denoted s_h^{SCC}, decreases in the bonus when $\theta r\sigma^2 < 1$ and increases in the bonus when $\theta r\sigma^2 > 1$. For ease of later comparison, I provide the derivative to b_h:

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

Finally, the derivation of the altruistic worker's PC (PCH) proceeds along the same lines:

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

This expression is not particularly insightful, but by taking the derivative to b_h it can easily be seen that compared to SCC, PCH is more inclined to slope downwards as long as $b_h \leq 1$:

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

This is clearly smaller than the derivative of SCC for all $b_h \leq 1$.

References

[1] Al-Ubaydli, Omar, Steffen Andersen, Uri Gneezy and John A. List (2007), "For Love or Money? Testing Non-Pecuniary and Pecuniary Incentive Schemes in a Field Experiment." Mimeo, University of Chicago.

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

[3] Anderhub, Vital, Simon Gächter and Manfred Königstein (2002), "Efficient Contracting and Fair Play in a Simple Principal-Agent Experiment." *Experimental Economics*, 5(1): 5-27.

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

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

[6] Bellemare, Charles and Bruce Shearer (2007), "Gift giving and worker productivity: Evidence from a firm-level experiment." *Games and Econonomic Behavior*, Forthcoming.

[7] Bellemare, Charles and Bruce Shearer (2009), "On the Relevance and Composition of Gifts within the Firm: Evidence from Field Experiments." IZA Working Paper No 4339.

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

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

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

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

[12] Dur, Robert, Arjan Non and Hein Roelfsema (2008), "Reciprocity and Incentive Pay in the Workplace". Tinbergen Institute Discussion Paper 2008-080/1.

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

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

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

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

[17] 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.

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

[19] 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.

[20] Frey, Bruno S. (1997), "Not Just for the Money: An Economic Theory of Personal Motivation" Cheltenham: Edward Elgar

[21] Gächter, Simon, Esther Kessler and Manfred Königstein (2009), "Performance Incentives and the Dynamics of Voluntary Cooperation", Discussion paper.

[22] Gneezy, Uri and John A. List (2006), "Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments." *Econometrica*, 74(5): 1365-1384.

[23] Güth, Werner, Wolfgang Klose, Manfred Königstein and Joachim Schwalbach (1998), "An Experimental Study of a Dynamic Principal-Agent Relationship." *Managerial and Decision Economics*, 19(4-5): 327-341.

[24] Hennig-Schmidt, Heike, Bettina Rockenbach, and Abdolkarim Sadrieh (2008), "In Search of Workers' Real Effort Reciprocity – A Field and a Laboratory Experiment." *Journal of the European Economic Association*, forthcoming.

[25] Kube, Sebastian, Michel André Maréchal and Clemens Puppe (2006), "Putting Reciprocity to Work – Positive versus Negative Responses in the Field." University of St. Gallen Department of Economics working paper series no. 2006-27.

[26] Kube, Sebastian, Michel André Maréchal, and Clemens Puppe (2008), "The Currency of Reciprocity – Gift-Exchange in the Workplace." Working paper, Institute for Empirical Research in Economics, University of Zurich.

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

[28] List, John A. (2006), "The Behavioralist Meets the Market: Measuring Social Preferences and Reputation Effects in Actual Transactions." *Journal of Political Economy*, 114(1): 1-37.

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

[30] Rigdon, Mary (2009), "Trust and Reciprocity in Incentive Contracting." *Journal of Economic Behavior & Organization,* 70(1-2): 93-105.

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

[32] 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.

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