CPB Discussion Paper

No 60
March 2006

Team incentives in public organisations

An experimental study

Jana Vyrastekova, Sander Onderstal and Pierre Koning

CPB Netherlands Bureau for Economic Policy Analysis
Van Stolkweg 14
P.O. Box 80510
2508 GM The Hague, the Netherlands

Telephone +31 70 338 33 80
Telefax +31 70 338 33 50
Internet www.cpb.nl

ISBN 90-5833-265-9

Abstract in English

Using a simple production game, we investigate whether public firms perform better when they increase the power of their workers' incentive schemes. In a laboratory experiment, subjects choose between a 'public firm' and a 'private firm' with team and individual incentives, respectively. When exposed to individual incentives, workers in the public firm increase effort in one parametrisation, but show a decrease in another. One reason for the latter observation is that reciprocators self-select in the public firm, rendering cooperation profitable.

Abstract in Dutch

In dit paper onderzoeken we middels een (laboratorium-)experiment de effectiviteit van prestatieprikkels bij publieke organisaties. Deelnemers aan het experiment kunnen kiezen tussen een 'publieke' en 'private' organisatie, met als verschil dat binnen de publieke organisatie teambeloning wordt toegepast; binnen de private organisatie daarentegen individuele prestatiebeloning. De introductie van individuele prestatiebeloning leidt bij de publieke organisatie tot gemengde resultaten. Individuele beloning heeft vooral kwalijke effecten wanneer deelnemers die van nature gemotiveerd zijn andere deelnemers te helpen daartoe ontmoedigd worden.

Contents

Summary

A restructuring of the public sector often involves proposals to increase the power of employees' incentive schemes. Recent examples of such proposals are with respect to teachers, employment offices, and medical practices. A priori, however, it is not clear whether high-powered incentives do improve performance under all circumstances. It is only if each worker's output is easy to measure with reasonable precision, and the interdependencies among workers' activities are low or undesirable that payments based on individual performance can be effective. In addition, the introduction of performance-based incentives may affect the composition of workers in public firms: better skilled workers may decide to join a firm that has switched to performance related pay, yet at the same time workers that are motivated to cooperate may decide to leave such firms. Still, the effect of self-selection is hardly explored in the literature. We intend to fill this gap.

In this paper, we investigate whether public firms perform better when they increase the power of their workers' incentive schemes. We do so using a laboratory experiment, allowing us to isolate the effect of the power of the incentive scheme on a worker's performance from other effects, particularly those related to self-selection of worker types to firms with high and low powered incentive schemes. In the experiment, subjects first choose between a 'public firm' and a 'private firm'. In these firms, workers perform two tasks in a simple production game. In Task 1, they produce their own output, while in Task 2, they increase the output of the colleague in their team. The firms pay their employees on the basis of both team performance and individual performance. The power of the incentive scheme increases if individual performance has a higher weight in a worker's payment, so that, in equilibrium, workers realise higher output. Team performance has a higher weight in the public firm than in the private firm, and hence the public firm has the lower-powered incentive scheme.

After 10 rounds of production in the firm they choose, all subjects participate in 10 more rounds and are paid according to the private firms payment scheme. A priori, it is not clear whether the performance of public firms will improve after the change of incentives. The higher power of the private firms' scheme may induce workers to exert more effort in Task 1. However, this effect may be counterbalanced because workers no longer have the incentive to reward their colleague for high effort in Task 1 with high effort in Task 2. In other words, reciprocal behaviour is not triggered by the private firm's incentive scheme because money-maximising workers exert high effort in Task 1 in the sub-game perfect Nash equilibrium. The latter effect may be strong because in the self-selection process, reciprocators may be more likely to choose to work for a public firm than money-maximisers. We let subjects submit strategies in the trust game to measure trust and reciprocity.

Our experimental findings show that increasing the power of the incentive scheme has ambiguous effects: workers in the public firm increase effort in one parameterisation, but show a decrease in another. In both parameterisations, we mainly observe high effort in Task 1 and low

effort in Task 2 under the private firm's payment scheme, while the opposite is true for the public firm. In the public firm, however, we find that a substantial fraction of the subjects opt for high effort in both tasks. The payment scheme's effect on total effort depends on the parameterisation. In the second parameterisation, where worker interdependencies are relatively strong, we find the total effort to be higher in de public firm. This suggests that reciprocal behaviour can be triggered and is beneficial when workers have to rely on the effort of co-workers. Indeed, the possibility to select the incentive scheme improves the economic performance of the public firm in the second parameterisation. We also observe that trust and reciprocity explain subjects' behaviour. The more a subject trusts and the higher her propensity to reciprocate: (1) the more likely it is that she will choose the public firm and (2) the more effort she will exert in Task 2.

The policy implications of our experiment are as follows: A public firm may or may not perform better if the power of workers' incentive schemes is increased. If the interdependency between workers is high and difficult to observe, a higher-powered incentive scheme may imply worse outcomes. For example, one may think of case managers at employment services, who often have to rely on their mutual expertise to find suitable jobs for their clients. Similar interdependencies may exist between medical doctors, or police officers. In contrast, when the gain from the unobservable tasks is low, then introducing individualistic incentives is the preferred choice. Another policy option is to strengthen the advantages of team pay, that is, develop sorting mechanisms to attract reciprocal workers, or strengthen signalling mechanisms within the organisation.

1 Introduction

A restructuring of the public sector often involves proposals to increase the power of employees' incentive schemes (Dixit (2002)). Recent examples of such proposals are with respect to teachers, employment offices, and medical practices (see Burgess and Ratto (2003) for a survey). A priori, however, it is not clear whether high-powered incentives do improve performance. As Gibbons (1998) argues in his review, high-powered incentives certainly have an impact on a worker's performance. However, he acknowledges that employers get 'what they pay for' : it is only if each worker's output is easy to measure with reasonable precision, and the interdependencies among workers' activities are low or undesirable that payments based on individual performance can be effective (see also Lazear (2000a)). This is precisely the problem in the public sector because public firms face measurability problems and complex multi-tasking (Besley and Ghatak (2003)), which often 'make inappropriate the naive application of magic bullet solutions like competition or performance-based incentives' (Dixit (2002), p. 697). In addition, the introduction of performance-based incentives may affect the composition of workers in public firms. As a result, better skilled workers may decide to join a firm that has switched to performance related pay, yet at the same time workers that are motivated to cooperate may decide to leave such firms. Still, the effect of self-selection is hardly explored in the literature.[1] We intend to fill this gap.

In this paper, we investigate whether public firms perform better when they increase the power of their workers' incentive schemes. We do so using a laboratory experiment in which subjects first choose between a 'public firm' and 'private firm' in a labour market. In the firms, workers perform two tasks in a simple production game. In Task 1, they produce their own output, while in Task 2, they increase the output of the colleague in their team. The firms pay their employees on the basis of both team performance and individual performance. We say that the power of the incentive scheme increases if individual performance has a higher weight in a worker's payment, so that in equilibrium, workers realise higher output. Team performance has a higher weight in the public firm than in the private firm, and hence the public firm has the lower-powered incentive scheme.

After 10 rounds of production in the firm they choose, all subjects participate in 10 more rounds and are paid according to the private firm's payment scheme. A priori, it is not clear whether the performance of public firms will improve after the change of incentives. The higher power of the private firms' scheme may induce workers to expend more effort in Task 1 because they obtain a higher payment for their own output. However, this effect may be counterbalanced because workers no longer have the incentive to reward their colleague for high effort in Task 1

[1] Notable exceptions are Lazear (2000a) (in a field study) and Keser and Montmarquette (2004) (in a laboratory experiment).

with high effort in Task 2. In other words, reciprocal behaviour is not triggered by the private firm's incentive scheme because money-maximising workers exert high effort in Task 1 in the subgame perfect Nash equilibrium. The latter effect may be strong because in the self-selection process, reciprocators may be more likely to choose to work for a public firm than money-maximisers. We let subjects submit strategies in the trust game to measure trust and reciprocity.

The reason for turning to experimental methodology is obvious when considering the measurement and identification problems that are connected to empirical data. Measurement problems may, for instance, arise in connection to workers' output and even if output is measurable, the researcher has no information on an individual's effort and preferences. Identification problems occur because it is not straightforward to isolate the effect of the power of the incentive scheme on a worker's performance from other effects (Prendergast (1999)). In contrast, in the laboratory we can observe the effort expended by the subject and measure her preferences, and we can expose the same individual to different payment schemes and observe her reaction, while keeping the rest of the environment constant.

Our experimental findings show that increasing the power of the incentive scheme has ambiguous effects: workers in the public firm increase effort in one parametrisation but show a decrease in another. In both parametrisations, we mainly observe high effort in Task 1 and low effort in Task 2 under the private firm's payment scheme. In the public firm, however, we find that a substantial fraction of the subjects opt for high effort in both tasks. The payment scheme's effect on total effort depends on the parametrisation. In the second parametrisation, where workers have better incentives to cooperate, we find the total effort to be higher in de public firm. This suggests that reciprocal behaviour can be triggered and can be beneficial when workers have to rely on the effort of co-workers. Indeed, the possibility to select the incentive scheme improves the economic performance of the public firm in the second parametrisation. We also observe that trust and reciprocity explain subjects' behaviour quite well. The more a subject trusts and the higher her propensity to reciprocate (1) the more likely it is that she will choose the public firm and (2) the more effort she will exert in Task 2. We develop a game-theoretic model to explain why trust and reciprocity influence subjects' behaviour.

The remainder of the paper is organised as follows. In Section 1.1, we discuss related literature. Section 2 contains the theoretical analysis of the production game. In section 3, we discuss the design of the experiment and our hypotheses. Section 4 includes the experimental observations. Section 5 concludes the paper.

1.1 Related literature

In this subsection we review some of the extensive literature on incentive schemes. Several empirical papers compare team and individual performance based payment schemes. Theory predicts that team incentives are low-powered because they will lead to free-rider problems: workers fail to internalise the benefits their effort has on others (Lazear (2000b)). Burgess et al. (2003) (employment service providers) and Vollaard (2003) (the police) confirm the presence of free-riding under team incentives. It seems, however, that the team is too broadly defined in these papers. For instance, in Burgess et al. (2003), teams comprise of up to 17 different and geographically distant offices in the UK.

Moreover, the literature offers mechanisms which explain why team incentives may work quite well in small teams. These mechanisms include mutual monitoring coupled with an easy information flow among team members and options for subjects to reciprocate among each other within a team (Kandel and Lazear (1992)). Indeed, Ittner and Pizzini (2005) (physicians), Lavy (2002) (teachers), and Knez and Simester (2001) (airlines) all provide evidence on teams in which team incentives do perform well.

Several experimental studies investigate individualistic and team-based payment schemes. Van Dijk et al. (2001) compare the outcomes of imposed individualistic, team, and tournament payment schemes in a real effort experiment. They find that the effort expended in teams is on average the same as that under individualistic pay, though the variance is higher. Tournaments elicit the highest effort and variance. Nalbantian and Schotter (1997) also compare payment schemes based on team and individualistic performance. They document that the team performance schemes perform much worse due to free riding, and the difference is even increasing over time.

To our knowledge, only Keser and Montmarquette (2004), like in our design, allow subjects to choose their payment scheme. They let the subject stay in the same two-player team for the duration of the experiment. By construction, the maximum payoff is higher under team incentives than under individualistic incentives. In this set-up, team incentives are popular, and frequently lead to high output. In our experiment, team incentives are a priori less attractive than in Keser and Montmarquette because in our design, the socially optimal payoff a subject can earn in the public firm is the same as in the private firm. Moreover, subjects are randomly rematched after each production game so that trigger strategies cannot explain the instances of cooperation we observe.

Several papers show that ‘social preferences’ are important in people’s decision making and may affect the success of incentive schemes. The experimental work by Ostrom et al. (1992) and Fehr and Gächter (2000) on social dilemma games lends considerable weight to the pervasiveness of reciprocal preferences among humans. Their findings support the observation that team incentives can be successful. Minkler (2002) discusses other reasons why

low-powered incentive schemes perform well, including ‘work moral’, intrinsic motivation, and positive reinforcement. In questionnaires, workers indicate that they very much appreciate non-pecuniary motivations like intrinsic incentives to perform the job (Deci (1971) and Frey and Jegen (2001)) as well as a fair relationship with the employer (Fehr and Schmidt (2004)). In a theoretical model, Francois (2000) shows that in the presence of ‘public service motivation’ , the government might prefer a public firm over a private firm to provide a public service, in order to elicit high effort from the motivated agents. In a similar setting, Delfgaauw and Dur (2004) show that public firms optimally attract motivated workers if effort is verifiable, and both motivated and ‘ lazy’ ones if it is not. Canton (2005) presents a multiple-task principal-agent model in which some agents are intrinsically motivated for the job they perform. He shows that high-powered incentives may not be optimal because they crowd-out intrinsic motivation.

Fehr and Schmidt (2004) study reciprocity in an experiment in which an employer offers a payment scheme to an employee. In their two-task production game, only the output of the first task is contractible, i.e. verifiable in court. The output of the second task is observable as well, but not verifiable. The employer can choose between a piece-rate contract (based on verifiable output only) and a contract which pays a fixed rate and an ex-post bonus if the employee works hard. The authors observe that employers choose the bonus contract much more frequently than the piece-rate contract, and that employees respond to this by putting a lot of effort into both tasks, and eliciting rewards (in the form of a bonus). Under the piece rate contract, the unrewarded task is neglected. Falk and Kosfeld (2005) find qualitatively the same results in a similar experiment. Our design is complementary to Fehr and Schmidt’s. Subjects in our experiment are not offered one contract (in a take-it-or-leave-it manner) but face a menu of contracts (like in the labour market). In addition, we address the reciprocity among employees rather than between employees and employers.

Finally, Holmstrom and Milgrom (1991) pioneer principal-agent models in which the agent can extend effort in several tasks. They find that high-powered incentives for one task can harm the principal’s objectives by destroying the agent’s incentives to make effort in another task that is not contractible. We address non-contractible tasks as well, namely cooperation between workers. In a similar way, Itoh (1991) analyses situations in which an employer prefers to provide team incentives (eliciting effort on an unobservable ‘helping’ task) as opposed to individualistic incentives on the observable task only. Our model is a simple version of this environment: we abstract from the stochastic relation between input and output in Itoh’s model.

2 Theory

We study a game in which individuals choose between a 'public firm'and ’private firm’ in a labour market. In both firms, two individuals form a team and play a simple production game. After describing the rules of the game, we study its properties, first under the assumption that individuals are only interested in their own monetary pay-off, and then that (part of) the population is reciprocal.

The rules of the production game are as follows. The two individuals (labeled 1 and 2) provide input in two subsequent tasks. In Task 1, each individual i independently chooses effort $e_i \in \{L,H\}$, where $0 \leq L < H$. In Task 2, after observing the effort of the other individual, each individual i chooses a reward $r_i \in \{0,R\}$, with $0 < R$. The higher effort and reward the higher the individual’s costs. Individual i pays $c(e_i) \leq e_i$, with $c(L) = 0 < c(H)$ and $k(r_i)$ where $k(0) = 0$ and $R > k(R) > 0$. An individual’s effort raises her own output, while her reward raises the output of her team member (which could be interpreted as helping the other person or being cooperative with her). More precisely, the relationship between the efforts and the rewards of players i and j forming one team and the output o_i of player i is the following

$$o_i = e_i + r_j$$

$\{i, j\} = \{1,2\}$. We assume that individual output is observable by the team manager, while the input is observable only by the team members and unverifiable vis-a-vis the court.

A parameter $\alpha \in \left[\frac{1}{2},1\right]$ characterises the environment in which the production team operates. Player i receives the following payment as a function of her output and the output of the other player:

$$p_i = \alpha o_i + (1-\alpha)\, o_j$$

$\{i, j\} = \{1,2\}$. In other words, α is a measure of the extent to which an individual’s output determines her payment. In the extreme case $\alpha = 1$, only her own output determines what she gets, while the other extreme $\alpha = \frac{1}{2}$ indicates that the payments are only based on total team performance. Note that there is a one-to-one relationship between the firm’s total output and the team’s total payment. In other words, more production is to the benefit of both the firm and the workers. Table 2.1 summarises the resulting two-stage game.

Proposition 1 characterises the subgame perfect Nash equilibrium (SPNE) for almost all values of $\alpha \in \left[\frac{1}{2},1\right]$.[2]

[2] We restrict our attention to situations in which a unique subgame perfect Nash equilibrium exists, which holds true for almost all values of α.. In all other cases, in at least one stage of the game, players are indifferent between the two actions, so that there is a larger range of subgame perfect Nash equilibria.

Table 2.1 Payoff matrices of the production game

Task 1	H	L
H	$H-c(H)$	$\alpha H+(1-\alpha)L-c(H)$
L	$\alpha L+(1-\alpha)H$	L
Task 2	R	0
R	$R-k(R)$	$(1-\alpha)R-k(R)$
0	αR	0

Proposition 1

For almost all values of $\alpha \in \left[\frac{1}{2},1\right]$*, the production game has a unique SPNE. In this equilibrium, both individuals choose effort* $e=H$ *[*$e=L$*] in the first task if*

$$\alpha > [<] \frac{c(H)}{H-L} \equiv \alpha_1.$$

In the second task, both choose reward $r=R$ *[*$r=0$*] if*

$$\alpha < [>] 1-\frac{k(R)}{R} \equiv \alpha_2.$$

This proposition has two important implications. First, the higher α the higher the output in the first task. Second, the higher α the lower the output in the second task. Observe that the players reach the Pareto optimum if and only if both play $e=H$ and $r=R$. In some parametrisations, this may be an equilibrium outcome for intermediate values of α.

We introduce some restrictions on the parameters. First, we assume that the following condition holds:

$$c(H) < H-L \tag{C1}$$

so that for a strictly positive mass of α's, effort $e=H$ can be obtained in a SPNE. Second, we wish to focus on the nontrivial case in which there is no α for which the Pareto efficient outcome can be achieved in a SPNE, that is:

$$\alpha_1 > \alpha_2. \tag{C2}$$

In the labour market, workers choose between a public firm and a private firm. The firms differ with respect to their payment scheme. We let INDI [TEAM] denote the game played in the private [public] firm, because in the private firm's payment scheme has a larger weight on individual [team] performance than the public firm. More precisely, if α_{INDI} [α_{TEAM}] denotes the private [public] firm's α, we assume that

$$\alpha_2 < \alpha_{TEAM} < \alpha_1 < \alpha_{INDI}.$$

The following proposition implies that a manager who in interested in high output prefers to confront the workers with INDI.

Proposition 2

In equilibrium, in INDI [TEAM], both individuals in the production team choose effort $e = H$ [$e = L$] and reward $r = 0$. Consequently, in INDI, higher total output is realised and hence, all players choose INDI.

Proof

Follows immediately from Proposition 1.

However, TEAM may yield a better outcome than INDI if the population contains sufficiently many reciprocal individuals, i.e. players who wish to cooperate as long as their team mate does so as well. A reciprocator plays the following 'tit-for-tat' strategy in TEAM: she starts off by choosing effort $e = H$ in Task 1, and she continues to cooperate by submitting reward $r = R$ in Task 2 if and only if the other team member chooses the high effort $e = H$ in Task 1 as well.

In order to make things more precise, let us assume that each individual has a value v for cooperation, i.e., she obtains additional utility v if in TEAM, both she and her team mate play $e = H$ in the first task and $r = R$ in the second. Moreover, it is common knowledge that individuals draw their value v independently from the same distribution function F which does not have strictly positive mass outside the interval $[\underline{v}, \overline{v}]$. In INDI, all choose $e = H$ in the first task, and $r = 0$ in the second, regardless of what the other player does. Reciprocation is not triggered in INDI because high effort in Task 1 is a dominant strategy and cannot be interpreted as an act of cooperation.

TEAM may have two types of equilibria. First, if $\overline{v}$ is sufficiently high, and F contains sufficient mass close to $\overline{v}$, all types above [below] a threshold value $\hat{v}$ play $e = H$ and $r = R$ [$r = 0$]. Second, if F is concentrated in the neighbourhood of 0, all players play $e = L$ and $r = 0$. The following proposition makes both claims more precise:

Proposition 3

Suppose the following inequality has a solution in the interval $[\underline{v}, \overline{v}]$ and that $\hat{v}$ is the smallest solution:

$$v(1 - F(v)) \geq k(R) - R(1 - \alpha_{TEAM}). \tag{2.1}$$

If

$$\alpha_{TEAM}(H - L + R(1 - F(\hat{v}))) \geq c(H) \tag{2.2}$$

then in the SPNE of TEAM, all types $v \geq \hat{v}$ [$v < \hat{v}$] play $e = H$ and $r = R$ [$r = 0$]. The firm's output is higher in TEAM than in INDI. Moreover, if (2.1) has no solution, all choose $e = L$ and $r = 0$. In that case, the firm's output is lower in TEAM than in INDI.

Proof

Because $\hat{v}$ is the smallest v that prefers to choose $r = R$ in Task 2, (2.1) follows from

$$(R+v)(1-F(v))+F(v)(1-\alpha_{TEAM})R-k(R) \geq \alpha_{TEAM}R(1-F(v)). \tag{2.3}$$

The LHS [RHS] of (2.3) refers to the expected payoff for type v if she chooses $r = R$ [$r = 0$], given that all types below v choose $r = 0$ and all types above v play $r = R$. Because $\hat{v}$ is the smallest solution, all types above [below] v strictly prefer to play $r = R$ [$r = 0$]. Equation (2.2) indicates that all prefer $e = H$ over $e = L$ in Task 1, and follows from

$$\alpha_{TEAM}R(1-F(\hat{v}))+H-c(H) \geq \alpha_{TEAM}L+(1-\alpha_{TEAM})H. \tag{2.4}$$

The LHS [RHS] of (2.4) refers to a player's expected pay-off if she choose $e = H$ [$e = L$] in Task 1 and $r = 0$ in Task 2 if her opponent plays the equilibrium strategy.

The equilibrium described in Proposition 3 states that high effort in the first task of TEAM is feasible in equilibrium under assumption (C2) and the following two additional conditions. First, there is a type $\hat{v}$ that weakly prefers to play $r = R$ rather than $r = 0$ if she and her team mate played $e = H$ in the first task and if all types $v \geq \hat{v}$ [$v < \hat{v}$] choose $e = H$ and $r = R$ [$r = 0$]. Second, none of the players prefers to play $e = L$ in the first task.[3]

Now, suppose that individuals can choose between INDI and TEAM. Let us assume that each individual i has an idiosyncratic belief $\hat{F}_i$ about the distribution of v with support $[\underline{v}, \overline{v}]$, and she believes that all other individuals share this believe. Now, suppose the population only consists of individuals with a low value v who believe that everybody else has a low value v as well. Then Proposition 3 implies that all prefer to play INDI rather than TEAM, as this would yield them $H - c(H)$ instead of L. However, if a player's belief $\hat{F}_i$ assigns sufficient mass to high values of v, then she has an incentive to enter TEAM as, according to Proposition 4, she expects that this scheme would give her a higher payoff relative to INDI. The following proposition makes these claims more precise.

Proposition 4

Suppose the following inequality has a solution in the interval $[\underline{v}, \overline{v}]$ and that $\hat{v}$ is the smallest solution:

$$v(1-\hat{F}_i(v)) \geq k(R)-R(1-\alpha_{TEAM}). \tag{2.5}$$

Moreover, assume that

$$\alpha_{TEAM}\left(H-L+R(1-\hat{F}_i(\hat{v}))\right) \geq c(H) \tag{2.6}$$

Then individual i prefers TEAM over INDI. Her preference is reversed if (2.5) has no solution.

[3] A sufficient condition for equation (2.1) having a solution is $v_{med} > 2k(R) - 2(1-\alpha_{TEAM})R$, where v_{med} is the median value, i.e., $F(v_{med}) = \frac{1}{2}$. It has no solution in the case that all players have value $v = 0$.

Proof

If the three conditions are satisfied, according to Proposition 3, individual i's utility in TEAM is at least $H - c(H) + \alpha R(1 - \hat{F}_i(\hat{v}))$, which is strictly larger than $H - c(H)$, the utility she would realise in INDI. In contrast, if (2.5) has no solution, Proposition 3 indicates that individual i's utility in TEAM is L, which is strictly smaller than $H - c(H)$ by condition (C1).

The above propositions have several implications. First, if the idiosyncratic beliefs about the distribution of others' value to reciprocate vary substantially, individuals may enter in both INDI and TEAM. Second, individuals who believe that sufficiently many others will reciprocate choose to enter TEAM and play $e = H$ in Task 1. Those who have a sufficiently high value of v will also choose $r = R$, the others will choose $r = 0$ in Task 2. Third, a firm currently using TEAM incentives can only increase the output by shifting incentives to the INDI scheme if the conditions of Proposition 3 are not satisfied. If they are, the team will produce less output in INDI than in TEAM.

3 The experiment

In this section, we describe the design of our experiment and the hypotheses that we wish to test based on the results from the theory.

3.1 Design

In the year 2004, we ran 9 experimental sessions using two parametrisations of the production game (see Tables 3.1 and 3.2). Altogether, 172 students from Tilburg University participated in the experiment; 166 of them continued into the production game (the remaining students were dropped when an odd number of them entered either of the two schemes). Participants were paid for all points they earned in the experiment (on average 13 Euro including a 5 Euro participation fee for a session lasting approximately 1.5 hours). The experiments took place in English. The experiments were fully computerised, programmed, and conducted using z-Tree (Fischbacher (1999)). Upon arrival at a session, participants were randomly seated at computer cubicles which were separated by blinds. During the experiment, communication other than via computer was prohibited.

Of the 172 subjects, 134 participated in the main design (described below), while 38 entered control sessions (see further below). In each experimental session of the main design, subjects had to make decisions at four different stages:

1. The trust game
2. A ‘labour market’ in which subjects chose between TEAM and INDI
3. The production game (10 rounds) in the chosen incentive scheme TEAM or INDI
4. The production game (10 rounds) in INDI.

Subjects only received instructions for the stage that they were on, and were not informed about the stages to follow. Moreover, so as not to contaminate further decision making in the main part of the experiment, no feedback on the trust game was given before the end of the experiment.

Let us discuss the experiment in detail. We use the trust game to measure subjects’ reciprocity and trust (i.e., their belief in others’ reciprocity). Berg et al. (1995) designed the trust game to mimic a situation in which two players, a sender and a receiver, can profit if trust exists between them. The sender has to decide how much of her 10 point endowment to transfer to the receiver. This money is then tripled and the receiver has to decide how much money (if any) to return to the sender. In the unique SPNE of the trust game, a money maximising receiver will return zero, so that a money-maximising sender will transfer zero. However, senders who expect receivers to be sufficiently reciprocal have the incentive to transfer a strictly positive amount of money. We, therefore, use a subject’s action in the role of sender to evaluate her level of trust, and her strategy in the role of receiver to measure her level of reciprocity.

In order to obtain a measure of both trust and reciprocity for each subject, we use a strategy method behind the veil of ignorance.[4] We asked subjects to submit strategies for both roles of the trust game, i.e., they first decided how much to transfer to the receiver in the role of sender, and then how much to return to the sender for every level of transfer the sender could make. At the very end of the experimental session, we let the computer decide at random which role each subject would play and to whom she would be matched. We paid them according to the strategies they submitted for the role that was assigned by the computer. A measure of a subject's trust is how much she transfers in the role of sender (which is an integer between 0 and 10).[5] A measure of a subject's reciprocity is the average fraction she returns in the role of receiver (which is a number between 0 and 1).

In the second stage of the experiment, we present subjects with a choice between the production games INDI and TEAM. Four sessions used parametrisation P1 and three sessions parametrisation P2, see Table 3.1. These parametrisations result in the pay-off matrices depicted in Table 3.2. The intuition behind the choice of parametrisations P1 and P2 is that in P1 subjects have more reason to deviate from the cooperative outcome than in P2. As a result, individuals will be less likely to be triggered to show reciprocal behaviour in P1 than in P2. If an odd number of subjects chose either scheme, one or two subjects were randomly excluded from continuing in the experiment, so that we could match the subjects into pairs. In each round, a subject was assigned to another anonymous co-player from among the subjects who chose the same scheme. This matching procedure was known to the subjects.

After subjects participated in these first 10 rounds, we informed them that they would play 10 more rounds. This time, however, we paid all of the subjects according to INDI. We rematched them after each round, but only among those who chose the same scheme. Having also finished these 10 rounds, the subjects learned the results of the trust game and their accumulated earnings for the whole experiment, and were paid by us in cash.

In order to evaluate the effect of sorting, we ran a control session for each parametrisation, with 18 and 20 subjects in P1 and P2 respectively. In both control sessions, the procedures were kept as similar to the other sessions as possible. Subjects submitted the trust game strategies, and then played 10 rounds of the production game under TEAM incentives in the given parametrisation. Afterwards, we informed subjects that they would participate in 10 more rounds and we exposed them to the TEAM scheme in the other parametrisation. We chose this approach so as to let subjects earn approximately the same amount of income as in the other sessions. In our analysis, we only compare TEAM without self-selection in rounds 1 to 10 to TEAM with

[4] Vyrastekova and Onderstal (2005) discuss this design and compare it to the standard trust game design.

[5] We are aware of the fact that the sender's motivation to transfer money in the trust game may go beyond the belief in positive reciprocity (Charness (2004)). For instance, risk and betrayal aversion might affect sender's decision to send money in the trust game (see Schechter (2005) and Bohnet and Zeckhauser (2004) respectively). However, there is evidence in the literature that using transfers as a measure of belief in reciprocity is reasonable (see Vyrastekova and Garikipati (2005)).

Table 3.1 Experiment parametrisations of the production games

	Parametrisation P1	Parametrisation P2
Number of sessions	4	3
Number of subjects	78	56
H	18	14
L	2	0
$c(H)$	12	8
R	16	14
$k(R)$	8	8
α_{INDI}	$\frac{7}{8}$	$\frac{6}{7}$
α_{TEAM}	$\frac{5}{8}$	$\frac{1}{2}$

Table 3.2 Experiment payoff matrices of the production games

Parametrisation P1

INDI			**TEAM**		
Task 1	H	L	Task 1	H	L
H	6	4	H	6	0
L	4	2	L	8	2
Task 2	R	0	Task 2	R	0
R	8	− 6	R	8	− 2
0	14	0	0	10	0

Parametrisation P2

INDI			**TEAM**		
Task 1	H	L	Task 1	H	L
H	6	4	H	6	− 1
L	2	0	L	7	0
Task 2	R	0	Task 2	R	0
R	6	− 6	R	6	− 1
0	12	0	0	7	0

self-selection (rounds 1 to 10 as well).

3.2 Hypotheses

The theoretical results of Section 2 provide a set of testable hypotheses concerning the firm subjects choose and their behaviour in these firms. In this subsection, we formulate the three hypotheses that we test using experimental data. Under the null hypotheses, we maintain the assumption of money-maximising subjects, while under the alternatives, we assume (beliefs in) reciprocal preferences.

Hypothesis 1: Self-selection

(H0) All subjects select INDI and in this scheme, $e = H$ in Task 1 and $r = 0$ in Task 2. (HA) Subjects who select INDI, opt for $e = H$ in Task 1 and $r = 0$ in Task 2. Subjects with high trust select TEAM and $e = H$ in Task 1. Those who (1) select TEAM, (2) are sufficiently reciprocal, and (3) observe history H,H in Task 1 choose $r = R$ in Task 2, the others who select TEAM choose $r = 0$ in Task 2.

As a consequence of the alternative in hypothesis 1, subjects entering TEAM will end up earning higher average profits than subjects entering INDI. In TEAM, two reciprocal subjects together earn $2R > 0$, a reciprocal and a nonreciprocal individual will earn $R - k(R) > 0$, and two nonreciprocal individuals will earn 0. This is more than in INDI, where subjects always earn 0. Self-selection into TEAM is less likely in P1 than in P2 because in P1, the critical conditions (2.5) and (2.6) on a subject's belief $\hat{F}_i$ to choose TEAM are stricter: $v(1-\hat{F}_i(v)) \geq 2$ and $\hat{F}_i(\hat{v}) \leq \frac{4}{5}$ versus $v(1-\hat{F}_i(v)) \geq 1$ and $\hat{F}_i(\hat{v}) \leq \frac{6}{7}$.

Next, we spell out the hypothesis most relevant for policy: does a team produce more the higher the power of the incentive scheme?

Hypothesis 2: The power of incentives

(H0) Subjects who selected TEAM produce more in INDI than in TEAM. (HA) Those who selected TEAM produce less in INDI than in TEAM.

Finally, we address the question, to which extent self-selection contributes to TEAM outperforming INDI - if it does at all. Subjects may select TEAM to express their willingness to provide high effort without explicit monetary incentives. We refer to this aspect as sorting. It is especially relevant when we think of public firms operating under team-based incentives. The incentive scheme is usually known to the workers entering the firm beforehand, i.e. it is one of the factors upon which they select the firm. If sorting matters then it generates a reason for the public firm to stick to TEAM incentives. We can evaluate the role of sorting for the success of team-based incentives by comparing the behaviour of subjects in TEAM who self-selected this scheme, to subjects who were forced to play TEAM. If more cooperation is found in the former group, then sorting is responsible for at least part of the success of TEAM.

Hypothesis 3: Sorting

(H0) Subjects who select TEAM when they can choose between TEAM and INDI produce the same output as those who are forced to play TEAM. (HA) Subjects who select TEAM when they can choose between TEAM and INDI produce more output than those who are forced to play TEAM.

Hypothesis 3 points out that initial sorting of subjects into TEAM and INDI may result in higher payoffs for subjects in TEAM. The driving force may be that those who choose TEAM trust more and are more reciprocal than those choosing INDI. If all subjects are forced to play TEAM, less reciprocity may be realised, so that the output in TEAM ends up being lower than it would be under sorting. We use the data from our control sessions to evaluate hypothesis 3.

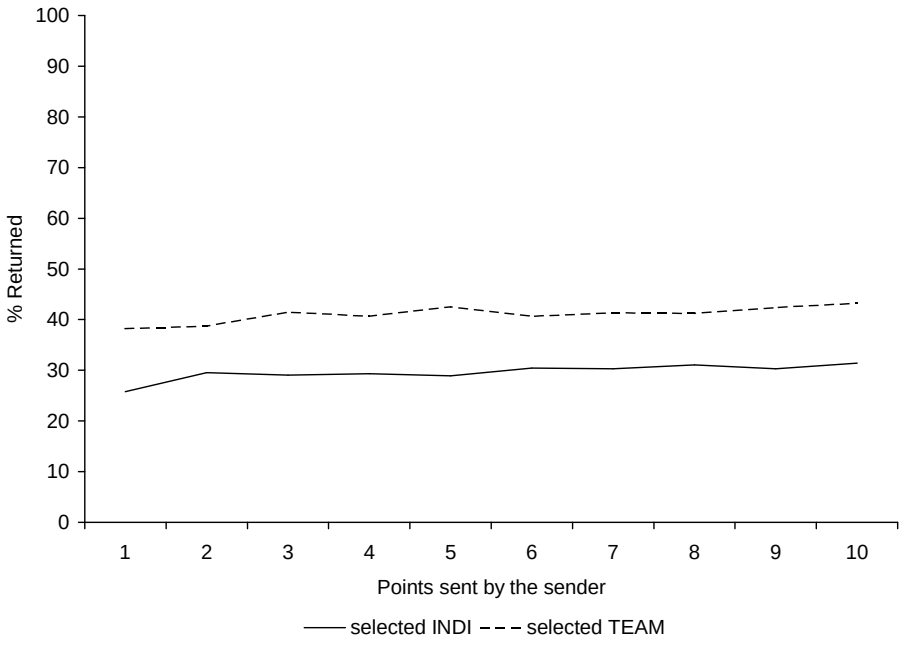
4 Data analysis

In this section, we address the three hypotheses formulated above. We reject the first null hypothesis for parametrisation P1, and all null hypotheses for P2, but not always in favour of the alternative hypothesis. Before discussing these observations in more detail, we wish to note that the findings in our trust game set-up do not differ substantially from the observations in Berg et al. (1995) (quoted in parentheses). In our experiment, senders tranfer, on average, 51% (52%) of their endowment to receivers, while 9% (6%) of them send nothing. The average return by receiver, as a fraction of sender's expenditure is 1.08 (0.89).

4.1 Self-selection

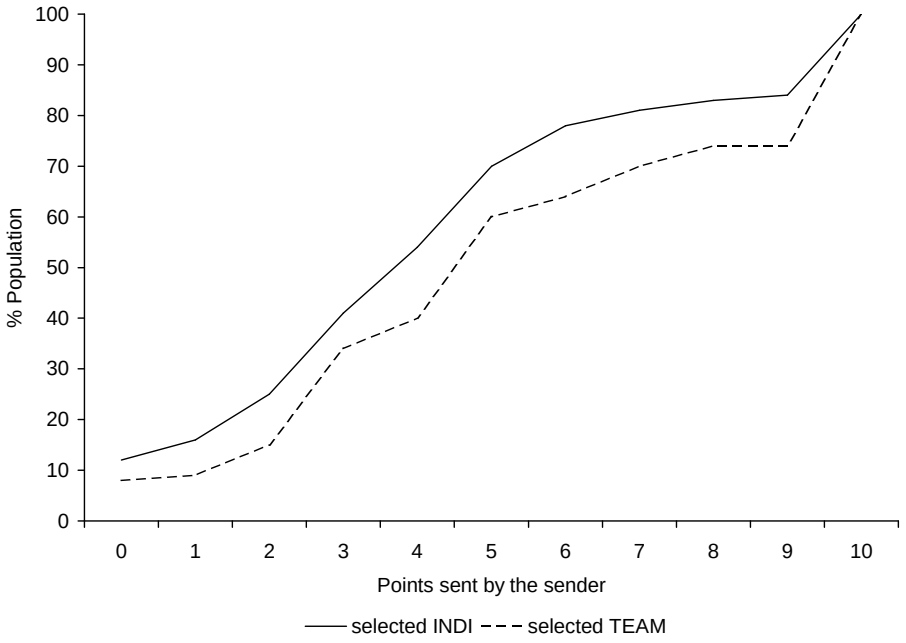
A substantial fraction of subjects chooses TEAM in both parametrisations: 34% in P1 and 44% in P2. This finding allows us to reject the null hypothesis that all select INDI. The fact that more subjects choose TEAM in parametrisation P2 is consistent with the prediction of Proposition 3, although the difference is not statistically significant. Figures 4.1 and 4.2 present subjects' choices in the trust game conditional on the firm they selected the labour market.

Figure 4.1 Reciprocity by subjects who selected INDI and TEAM.



From Figures 4.1 and 4.2, it becomes clear that those who choose TEAM have both more trust and higher reciprocity. Subjects selecting INDI send on average 4.6 points while those selecting TEAM send significantly more, 5.5 points (one-sided Mann-Whitney U test, $p = 0.049$). Subjects selecting TEAM return a fraction that leaves senders' investment profitable (i.e., more

Figure 4.2 Cumulative distribution of 'trust' by subjects who selected INDI and TEAM



than one third of the received number of points), unlike subjects selecting INDI who return less than senders sent to them.

Table 4.1 includes Probit estimates for the choice of TEAM. The outcomes show that there is a non-linear relationship between trust and reciprocity on one hand, and the choice of the firm on the other. Both trust and reciprocity have a positive effect on the choice of TEAM (as we have just observed using non-parametric tests). However, the interaction term is negative and (weakly) significant. Its parameter estimate implies that given values of trust [reciprocity] above 7 [0.57], higher reciprocity [trust] makes entry into TEAM less likely.

Table 4.1 Probit estimates of choice of TEAM. The coefficients are expressed as marginal effects.

Variable	Coefficient estimate	Standard error
Trust	0.057*	0.024
Reciprocity	0.740*	0.330
Trust * reciprocity	− 0.100**	0.052
P2-dummy	− 0.059	0.089
# observations	134	
Log Likelihood	− 85.44	

* indicates significance at less than 5%.
** indicates significance at less than 10%.

This observation has two implications. First, we reject null hypothesis 1 in favour of its alternative: more trust makes entry into TEAM more probable. Second, TEAM may attract 'free-riders': people who believe that sufficiently many other are reciprocators, without having the attention to reciprocate themselves. In other words, these subjects may imitate behaviour of reciprocators in Task 1, but free-ride on them in Task 2. Consequently, TEAM scheme attracts reciprocal as well as money-maximising subjects, as long as their trust is sufficiently high.

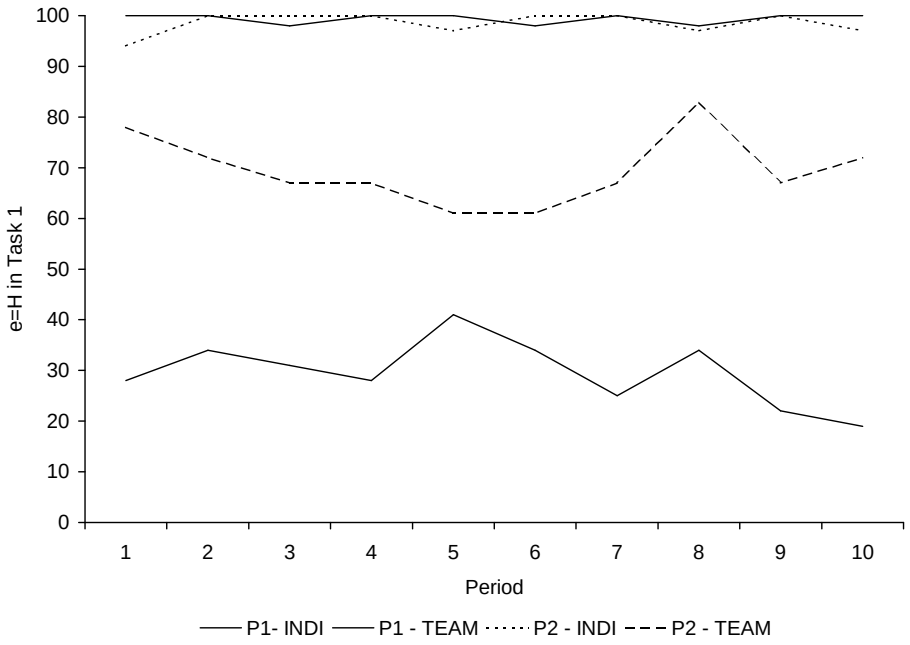
Observe that the P2-dummy (that equals 1 if subject participates in parametrisation P2) is not significantly different from zero. This implies that the hypothesis is rejected that subjects in P1 are more likely to choose INDI than those in P2.

Observation 1a (Self-selection):

A nonnegligible fraction of subjects (more than one third in either of the two parametrisations) selects TEAM. On average, subjects who do so trust more and are more reciprocal than subjects who opt for INDI. However, money-maximisers also enter TEAM, generating free-rider problems.

We now turn to analysing subjects' strategies in the firm of their choice. Figure 4.3 depicts their behaviour in Task 1.

Figure 4.3 Action $e = H$ in Task 1 for P1 and P2



From Figure 4.3, we derive that in Task 1, subjects in INDI choose nearly exclusively $e = H$ (which is a dominant strategy). Also in TEAM, we observe $e = H$: in parametrisation P1 [P2] on average 31% [69%] of subjects choose $e = H$. In the light of Proposition 3, this observation is

somewhat surprising: we expect someone who enters TEAM to always play $e = H$, hoping that her team mate will choose $r = R$ in Task 2. One explanation for this discrepancy is that subjects entered TEAM by mistake. There is indeed support for this suggestion. Subjects choosing $e = L$ in TEAM send significantly less points in the trust game than subjects choosing $e = H$ in the same scheme (5.2 points vs. 6.1 points in P1 and 4.1 points vs. 6.4 points in P2 with p-values of the Mann-Whitney U tests being $p = 0.059$ and $p = 0.000$, respectively). This means that their trust is lower than for subjects entering TEAM who choose $e = H$. However, low trust implies that they should prefer INDI (see Proposition 3).

Figure 4.4 Action $r = R$ in Task 2 for P1 and P2

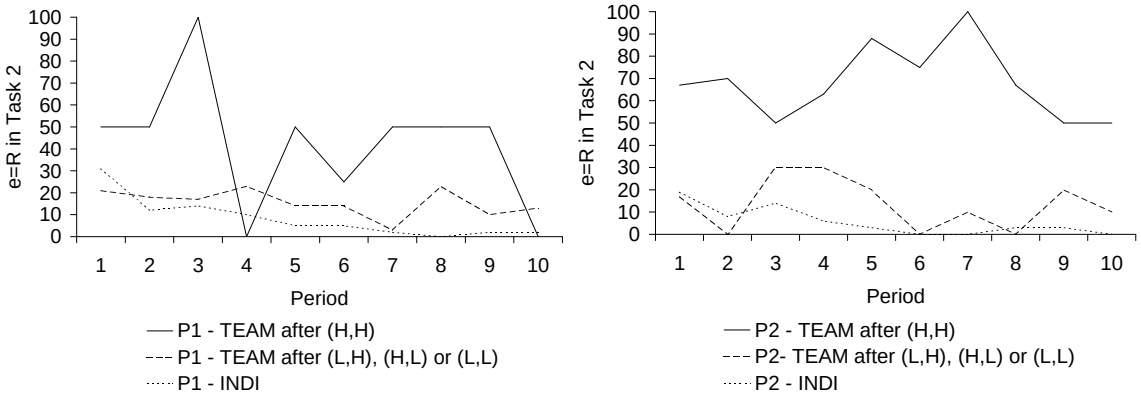


Figure 4.4 indicates that high effort in Task 2 is much more likely in TEAM than in INDI, especially if both subjects gave high effort in Task 1. We use a conditional logit model to further investigate this observation (see Table 4.2 for the estimationresults). This specification accounts for subject-specific effects that are likely to be correlated and to affect the behaviour of a subject in both Tasks. For instance, reciprocal subjects might be more likely to believe that others are reciprocal, so that they choose high effort in the Task 1, and these might be the same subjects who are more inclined to reciprocate in the Task 2. Indeed, we observe that those who choose $r = R$ in TEAM are more likely to be reciprocators than those who choose $r = 0$. The former return on average 51% [39%] of the received amount to the sender in P1 [P2], while the latter return on average 39% [25%]. The difference is significant in both cases (Mann–Whitney U test, $p = 0.090$ [$p = 0.008$]). Thus, without any controls, we would overestimate the coefficient describing the effect of high effort on the probability of reciprocation. For both P1 and P2 parametrisations, it is apparent that the probability of choosing $r = R$ is highest when both players chose $e = H$ in Task 1 (the coefficients on any other history observed is significant and negative), supporting the equilibrium derived in Proposition 3.

Table 4.2 **Conditional Logit model estimates for the probability of r=R in Stage 2 (standard errors between parentheses).**

Task 1 action	$e \in \{L,H\}$	Parametrisation	
Player	Co-player	P1	P2
H	H	reference group	reference group
H	L	− 1.54 (0.64)	− 3.41 (0.85)
L	H	− 1.63 (0.68)	− 3.13 (0.89)
L	L	− 1.91 (0.61)	− 1.98 (0.96)
Number of observations		190	120
Dropped (no variation)		130	60
Log likelihood		− 70.809	− 32.249

Observation 1b (Self-selection):

Those who select TEAM and observe history H,H in Task 1 are more likely to choose $r = R$ in Task 2 than those who observe another history. For subjects who are more reciprocal, this effect is stronger.

4.2 The power of incentives

Is it economically profitable to increase the power of the incentive scheme? Table 4.3 summarises the economic performance of INDI and TEAM. It contains the average profits per subject and round in the self-selected scheme in both parametrisations. In P1 [P2], selecting INDI is significantly more [less] profitable than TEAM for both the firm and the workers. Moreover, in both parametrisations, individuals who self-select into TEAM earn significantly more than the payoff predicted by the Nash equilibrium for this scheme.

Table 4.3 **Average and Nash equilibrium payoffs per subject per round in the selfselected scheme (rounds 1 to 10) and MannWhitney U test (standard errors between parentheses).**

	Parametrisation P1	Parametrisation P2
Payoff in SPNE		
INDI	6	6
TEAM	2	0
Average payoff		
INDI	6.6 (0.20)	4.6 (0.28)
TEAM	6.3 (0.16)	6.6 (0.39)
Mann-Whitney U test INDI-TEAM	p=0.000	p=0.002

These observations suggest that the firm can increase output by switching from TEAM to INDI in parametrisation P1, but not in P2. We now report the impact of doing so on the behaviour of

the subjects in the last stage of the experiment, in which all played INDI. Figures 4.5 and 4.6 display actions chosen in rounds 11 to 20 in Task 1 and Task 2, respectively. They correspond closely to the SPNE.

Figure 4.5 Action $e = H$ in Task 1 for P1 and P2

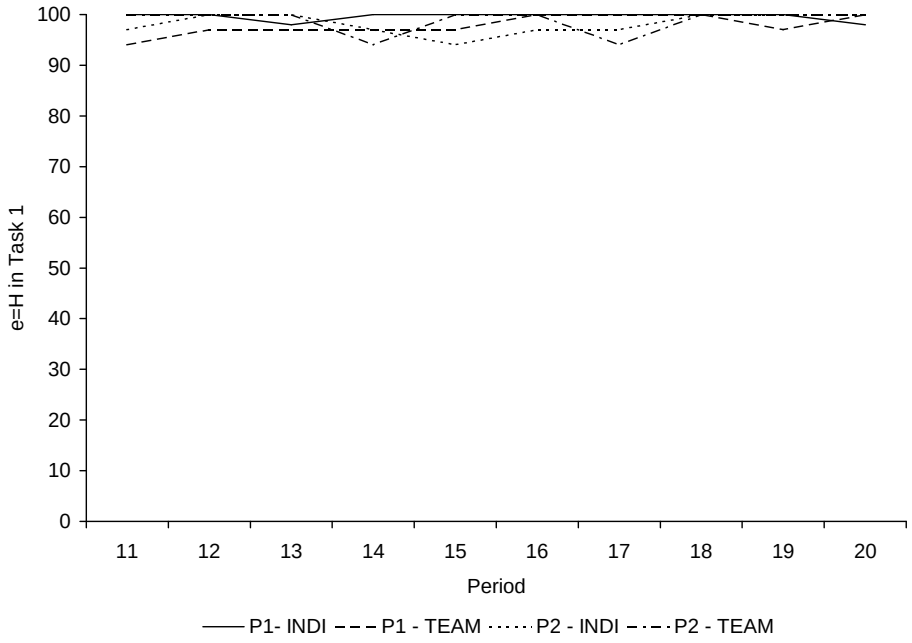


Figure 4.6 Action $r = R$ in Task 2 for P1 and P2

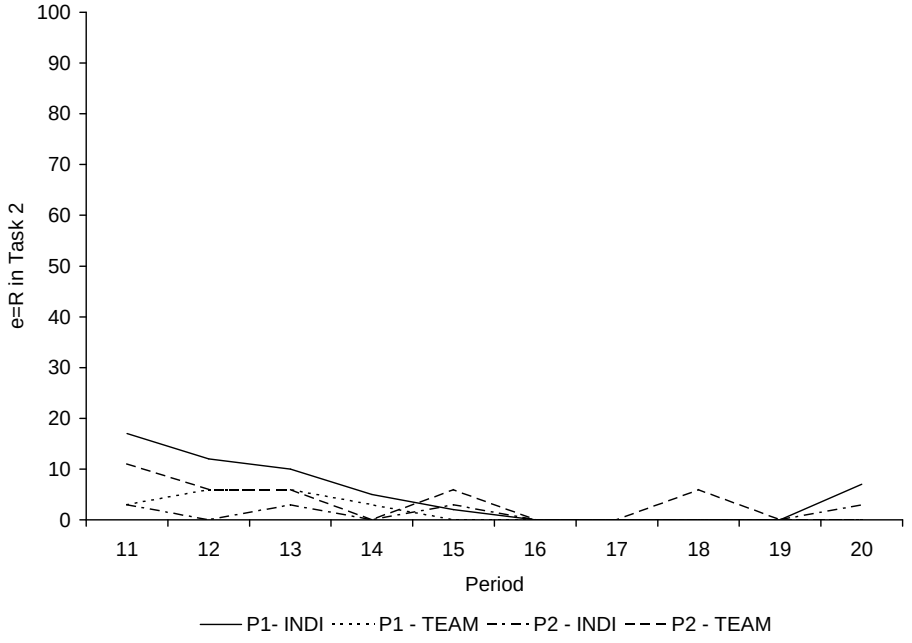


Table 4.4 presents the average payoffs for those who chose TEAM in both TEAM (periods 1 to 10) and in INDI (periods 11 to 20). The economic performance of the subjects who chose TEAM significantly changes with they are forced to play INDI. However, the change is not unidirectional. In parametrisation P1, we observe a significant increase in earnings (although the average earnings still fall short of the Nash equilibrium prediction). The opposite is found in parametrisation P2. Note that these observations cannot be related to the subject pool composition, because subjects are informed and are matched in the same subset of the pool as when they play according to the self-selected scheme.

Table 4.4 Average and Nash equilibrium payoffs for 10 rounds in the forced payment scheme COMP only for subjects who previously self-selected into COOP.

	Parametrisation P1	Parametrisation P2
Payoff in SPNE		
INDI	6	6
TEAM	2	0
Average payoff		
TEAM (rounds 1-10)	4.6 (0.28)	6.1 (0.11)
INDI (round 11-20)	6.6 (0.39)	6.1 (0.18)
Wilcoxon Signed Ranks test	p=0.001	p=0.044

Observation 2 (Team vs. individual incentives):

Subjects who choose TEAM, when exposed to INDI, improve performance in parametrisation P1, but not in P2.

4.3 Sorting

We have observed that in P2, TEAM performs better than INDI. Is the success of TEAM explained by the fact that reciprocal subjects choose TEAM in the labour market? We compare the actions in the production game in Figure 7 (Task 1) and Figure 8 (Task 2) with the control group which was forced to play TEAM. We find that in P1, sorting has no effect ($p = 0.909$ Mann-Whitney U test). However, in parametrisation P2, subjects earn significantly more in the sessions when they sort themselves into TEAM than when we force them to do so ($p = 0.016$ Mann-Whitney U test). More specifically, those who self-select are significantly more likely to choose $e = H$ in Task 1 (Mann-Whitney U test, $p = 0.012$), and $r = R$ in Task 2 (Mann-Whitney U test, $p = 0.005$).[6]

[6] In P2, subjects who are forced to play TEAM perform worse than subject who choose INDI, in contrast to those who choose TEAM.

Figure 4.7 Action $e = H$ in Task 1 for parametrisations P1 and P2 in the self-selected TEAM and in control TEAM sessions

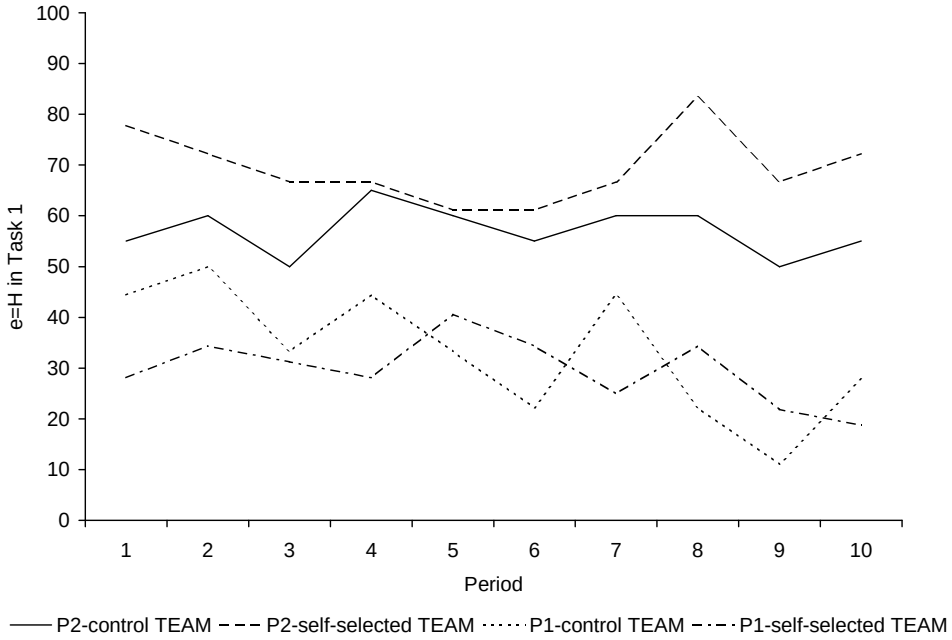
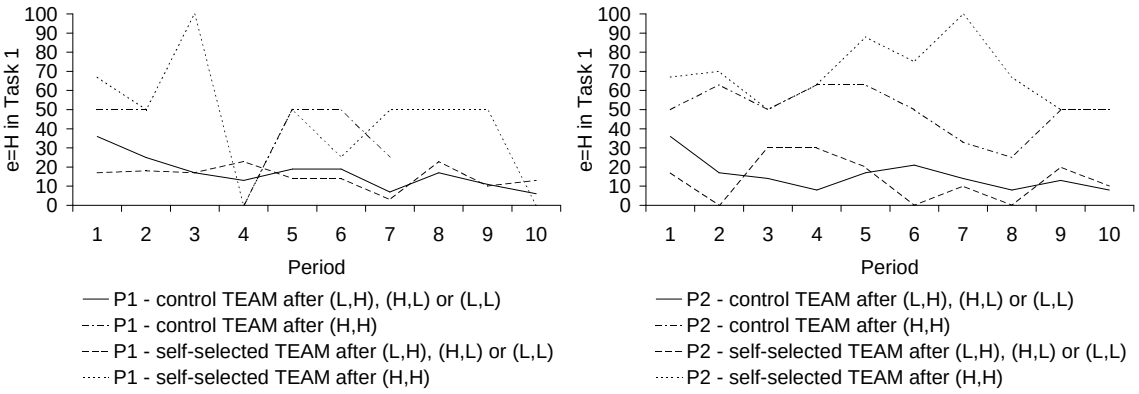


Figure 4.8 Action $r = R$ in Task 2 for parametrisation (i) P1 and (ii) P2 in the self-selected TEAM and in control TEAM sessions



Observation 3 (Sorting):

The impact of sorting on subjects' behaviour in TEAM is small in parametrisation P1. In contrast, in P2, those who are forced to play TEAM perform worse than those who select team when they have the opportunity to choose between INDI and TEAM. In other words, sorting partially explains why the increase in the power of the incentive scheme may have counterproductive effects.

5 Conclusions

The question of properly incentivising workers in public firms has been on the agenda for some time. High-powered incentives and proper measures of individual performance stand high on the list of policy instruments that should improve the performance of the public sector. In this paper, we have addressed the question of whether public firms perform better when they increase the power of their workers' incentive schemes. We have answered this question using a laboratory experiment. Our subjects could choose between a 'public' and a 'private' firm offering low-powered and high-powered incentives respectively. Afterwards, subjects played a two-task production game in the firm they selected. In Task 1, they produced their own output, while in Task 2, they could increase the output of the colleague in their team. We measured subjects' trust and reciprocity using an auxiliary trust game experiment.

In our experiment, a substantial fraction of subjects opted for the public firm, contrasting with the predictions of standard theory. Moreover, production increased under higher-powered incentives in parametrisation P1, but not in P2. We found two explanations for this observation. First, in P2, subjects have less reason to deviate from the cooperative outcome than in P1. As a result, reciprocal behaviour is more likely to be triggered in P2. Second, we observe that in P2, sorting explains why the public firm performs better under low-powered incentives than under high-powered incentives. When all subjects are forced to opt for the public firm, its output decreases relative to the situation in which subjects can choose between the public firm and the private firm.

The policy implications of our experiment are as follows: A public firm may or may not perform better if the power of workers' incentive schemes is increased. If the interdependency between workers is high and difficult to observe, and their incentive to cooperate is high in the case of a low-powered incentive scheme, a higher-powered incentive scheme may imply worse outcomes. For example, one may think of case managers at employment services, who often have to rely on their mutual expertise to find suitable jobs for their clients. Similar interdependencies may exist between medical doctors and police officers. In contrast, when the gain from the unobservable tasks is low, then introducing individualistic incentives is the preferred choice. One more policy option is to strengthen the advantages of team pay, that is, develop sorting mechanisms to attract reciprocal workers, or strengthen signalling mechanisms within the organisation.

According to our knowledge, our experiments are the first to disantangle the impact of workers' sorting themselves into firms with various incentive schemes and the role of signalling for the success of team-based incentives. Morever, by using the auxiliary trust game expeirment, we are able to link the performance of teams to trust and reciprocity, lending support to the theories relying on these explanations (e.g. Lazear (2000b)). Finally, we wish to note that we used a very cautious design. Our subjects did not have the opportunity to sort themselves into

teams endogenously or to build reputation. If subjects are able to form groups endogenously on the basis of historical information or repeated interaction, this could alleviate the free-rider problem. These options remain open for future research.[7]

[7] Keser and Montmarquette (2004) is a promising first step.

References

Berg, J., J. Dickhaut and K. McCabe, 1995, Trust, reciprocity and social history, *Games and Economic Behavior*, vol. 10, pp. 122–142.

Besley, T. and M. Ghatak, 2003, Incentives, choice, and accountability in the provision of public services, *Oxford Review of Economic Policy*, vol. 19, pp. 235–246.

Bohnet, I. and R. Zeckhauser, 2004, Trust, risk and betrayal, *Journal of Economic Behavior and Organization*, vol. 55, pp. 467–484.

Burgess, S., C. Propper, M. Ratto and E. Tominey, 2003, Incentives in the public sector: Some preliminary evidence from a uk government agency, Discussion paper.

Burgess, S. and M. Ratto, 2003, The role of incentives in the public sector: Issues and evidence, *Oxford Review of Economic Policy*, vol. 19, pp. 285–300.

Canton, E., 2005, Power of incentives in public organizations when employees are intrinsically motivated, *Journal of Institutional and Theoretical Economics*, vol. 161, no. 4.

Charness, H., 2004, Attirbution and reciprocity in an experimental labor market, *Journal of Labor Economics*, vol. 22, pp. 665–688.

Deci, E., 1971, Effects of externally mediated rewards on intrinsic motivation, *Journal of Personality and Social Psychology*, vol. 18, pp. 101–115.

Delfgaauw, J. and R. Dur, 2004, Incentives and workers' motivation in the public sector.

Dixit, A., 2002, Incentives and organizations in the public sector: An interpretative review, *Journal of Human Resources*, vol. 37, pp. 696–727.

Falk, A. and M. Kosfeld, 2005, The hidden cost of control.

Fehr, E. and S. Gächter, 2000, Cooperation and punishment in public goods experiments, *American Economic Review*, vol. 90, pp. 980–994.

Fehr, E. and K. Schmidt, 2004, Fairness and incentives in a multi-task principal-agent model, *Scandinavian Journal of Economics*, vol. 106, pp. 453–474.

Fischbacher, U., 1999, Z-tree: Zurich toolbox for readymade economics experiments.

Francois, P., 2000, 'public service motivation' as an argument for government provision, *Journal of Public Economics*, vol. 78, pp. 275–299.

Frey, B. and R. Jegen, 2001, Motivation crowding theory: A survey of empirical evidence, *Journal of Economic Surveys*, vol. 15, pp. 589–611.

Gibbons, R., 1998, Incentives in organizations, *Journal of Economic Perspectives*, vol. 12, pp. 115–132.

Holmstrom, B. and P. Milgrom, 1991, Multitask principal-agent analyses: Incentive contracts, asset ownership and job design, *Journal of Law, Economics and Organization*, vol. 7, pp. 24–52.

Itoh, H., 1991, Incentives to help in multi-agent situations, *Econometrica*, vol. 59, pp. 611–636.

Ittner, C. and M. Pizzini, 2005, Team-based compensation in professional service firms, manuscript.

Kandel, E. and E. Lazear, 1992, Peer pressure and partnerships, *Journal of Political Economy*, vol. 100, pp. 801–817.

Keser, C. and C. Montmarquette, 2004, Voluntary teaming and effort.

Knez, M. and D. Simester, 2001, Firm-wide incentives and mutual monitoring at continental airlines, *Journal of Labor Economics*, vol. 19, pp. 743–772.

Lavy, V., 2002, Evaluating the effect of teachers' group performance on pupil achievement, *Journal of Political Economy*, vol. 110, pp. 1286–1317.

Lazear, E., 2000a, Performance pay and productivity, *American Economic Review*, vol. 90, pp. 1346–1361.

Lazear, E., 2000b, The power of incentives, *American Economic Review*, vol. 90, pp. 410–414.

Minkler, L., 2002, Shirking and motivation in firms: Survey evidence on worker attitudes.

Nalbantian, H. and A. Schotter, 1997, Productivity under team incentives: An experimental study, *American Economic Review*, vol. 87, pp. 314–341.

Ostrom, E., R. Gardner and J. Walker, 1992, Covenants with and without sword: Self-governance is possible, *American Political Science Review*, vol. 86, pp. 404–417.

Prendergast, C., 1999, The provision of incentives in firms, *Journal of Economic Literature*, vol. 107, pp. 7–93.

Schechter, L., 2005, Traditional trust measurement and the risk confound: An experiment in rural paraguay, *Journal of Economic Behavior and Organization*, vol. forthcoming.

Van Dijk, F., J. Sonnemans and F. van Winden, 2001, Incentives systems in a real effort experiment, *European Economic Review*, vol. 45, pp. 187–214.

Vollaard, B., 2003, Performance contracts for police forces.

Vyrastekova, J. and S. Garikipati, 2005, Beliefs and trust: An experiment.

Vyrastekova, J. and S. Onderstal, 2005, The trust games behind the veil of ignorance: A note on gender differences.