Teachers' Performance Pay: External Incentives and Effort Allocations in Multi-tasking Problem*

Shraman Banerjee† Timothy C. Salmon‡

August 17, 2015

Abstract

Recent education policies have proposed replacing the fixed wage structure for teachers with merit pay based on students' test scores to improve teaching quality. One of the key claims of those who oppose the policy is that such incentives might induce teachers to teach only to the test, undermining development of other higher-order learning skills among students. Another claim by opponents is that there might be an effect which is best described as cross task crowding out of intrinsic incentives where financial incentives on one task diminish intrinsic incentives for engaging in another. To address these issues we study the effects of similar incentives on individual behaviors experimentally where subjects invest in two projects one of which generates more to their personal earnings while the other is more charitable. We also test to determine how adding risk to the incentives and allowing individuals to see the choices of others affects behavior. Our findings provide support for the claim that teachers might indeed shift towards teaching to the test but we do not find that evidence of any cross task crowding out of intrinsic incentives among the experimental subjects. We also find that providing social information does not increase the effort in the charitable dimension and that making the incentive pay risky does not decrease its ability to motivate subjects to provide effort along the individually incentivized dimension.

JEL Classification: C91, D03, D86, J33

Keywords: multi-tasking, performance pay, group norm, discouragement effect, experiment

*The authors specifically thank Danila Serra for providing helpful suggestions and feedback and Jing Li for her assistance during running the experimental sessions. The authors also wish to thank all the seminar participants at Southern Methodist University for useful discussions and comments. This research project was approved by the Institutional Review Board at Southern Methodist University.

†Southern Methodist University, Department of Economics, 3300 Dyer Street, Suite 301, Umphrey Lee Center, Dallas TX $75275 - 0496$. Email: sbanerjee@smu.edu

‡Southern Methodist University, Department of Economics, 3300 Dyer Street, Suite 301, Umphrey Lee Center, Dallas TX $75275 - 0496$. Email: tsalmon@smu.edu

1 Introduction

Those who educate children well are more to be honored than parents, for these only gave life, those the art of living well. – **Aristotle**

A recent debate that is widespread in the educational policy circles is: Should we introduce Performance-pay for teachers in schools as an incentive policy to improve the quality of teaching? The traditional wage structure for teachers was a fixed pay which would increase only on seniority or on tenure. On the other hand this performance-pay policy attempts to evaluate a teacher's quality and pay an increased wage based on their impact on students' standardized test scores (a measure called 'value added', introduced by Hanushek, 1971).[1] Quite a number of school districts in the US have already replaced this traditional wage structure by a merit based system or an incentive pay to elicit higher teaching efforts from teachers. Local and state officials are currently implementing performance-pay programs in states such as Alaska, Florida, Idaho, Iowa, Minnesota, New Mexico, North Carolina and Texas. Mark Wallace of the Center for Workforce Effectiveness, a consulting firm specializing in compensation, recently noted that nearly three out of every five states in US have enacted legislation requiring localities to explore performance pay (Delisio, 2003).

Although this performance-pay policy is in line with the standard practice in any organization of incentivizing employees to achieve a desired outcome and is the basic foundation of agency theory, the policy has been opposed in some of the studies in the education literature as well as by different teachers' unions in various parts of the US. This opposition stems primarily from two broad issues. The first major issue has been raised in the education literature. According to their claims the duties that a teacher performs inside a classroom is multidimensional (Hannaway, 1991). An incentive weighing heavily on test performances of the students might induce teachers to allocate their efforts more to teaching to the tests sacrificing some other aspects of teaching, such as generating curiosity and higher-order learning skills among the students which, they claim, forms an important component in the overall learning and education level of a society (Herman, 1992). Earl *et al* (2003) also claims that if the pay-for-performance policy evaluates teachers solely on the students' test scores, teachers might skew their efforts in the direction of activities that would increase the scores of the students in the standardized tests. So in spite of increasing students' test scores an incentive pay might affect adversely the overall learning and human capital formation of the society which would undermine the basic goal of these standardized test to improve the overall education level of the society. This is a potential issue, which the studies in education literature claim, has been ignored by the proponents of the incentive pay policy. The theoretical foundation of this multidimensionality of teachers' effort choices in a principal-agent framework was described by Holmstrom and Milgrom (1991) where incentives in one task steers away efforts from the other tasks where there are relatively less incentives. This is termed as the multi-tasking moral hazard problem in the literature.

Another potential issue though not mentioned in any formal education or economics literature, has been raised by several teachers' unions. These unions claim that teachers are already motivated to their duties of educating students and they need no incentives to perform their duties. Moreover

[1] A standardized test is one that is administered and scored under uniform and controlled conditions (Payne, 2003).

incentives might act as a source of demotivation for them. This claim can be subtly related to a branch of economics literature called motivation crowding out theory (Benabou and Tirole, 2000). This literature shows that if individuals have some intrinsic motivation to perform a task, providing external incentives might crowd out their intrinsic motivation for that task. Although all of this literature deals with motivation crowding out for a single task, what the issue of the teachers' unions translate into is the presence of some cross-crowding out of motivation for teachers. Performance-pay based on student test scores might crowd out teachers' motivations from the task of generating higher-order learning skills among students which are more socially beneficial.

To address these issues we conduct a laboratory experiment to study the effects of similar financial incentives on individual behaviors using a multi-tasking principal-agent framework. The objective of our paper is to answer two major questions. When individuals face two tasks one of which generates more to their personal earnings while the other is more socially beneficial, does the personal incentive on one task substitute away efforts from the other task ? This is related to the theoretical predictions of incentive effects of Holmstrom-Milgrom (1991) that under performance-pay teachers substitute efforts to teaching to the test which earns them bonus wage rather than helping students on other aspects of learning. The second question that our paper answers is whether external incentives on one task results in a cross-crowding out of their intrinsic motivation on the other task which is socially beneficial. The issue of cross-crowding out of motivation of the teachers has not been raised by the education literature and has only been mentioned by some of the teachers' unions. Our paper tests whether these claims are justified, since this would imply that the adverse effects of incentive pay on the overall learning of the society goes beyond what the theory predicts.

Past empirical literature on teachers' incentive pay studies the effects of this wage policy on students' test score gains. This paper looks into this issue differently as we study the effects of this policy on teachers' effort allocation as well as on their intrinsic motivation towards teaching. Effort choices and intrinsic motivation of the teachers are neither observable nor perfectly measurable from naturally occurring data. So the causal relationship between bonus pay and the teachers' effort choices can be empirically tested only imperfectly through students' test scores. So we use laboratory experiments to test this issue in a controlled environment. In lab we can actually observe how individuals allocate their resources (like time or effort) rather than focussing on just the outcome (which is the test scores). The controlled environment also helps us to abstract from other factors in the education environment affecting the teachers' choices, and focus primarily on the relationship between bonus and effort choices. These other factors may be the teaching environment in a particular school, the overall teaching infrastructure in the country as well as in a particular school, the education policy of the government etc.

Since a teacher's efforts towards teaching to the test are not perfectly related to the scores of the students, we can assume a teacher's wage to be a non-deterministic function of its efforts on teaching to the test. There are certain random factors involved in the test score generation of the students which are not in control of the teachers, such as students' inherent abilities, some random events on the test day etc. Thus a high effort on teaching to the test does not necessarily imply higher earning for the teachers if these outside random factors undermine the test scores of the students. This might discourage them to put high effort on teaching to the test and lead them to revert back to the other tasks of generating higher-order learning skills among students for which they have intrinsic motivation. Our paper tests whether this non-deterministic nature of earnings leads to a dynamic discouragement effect on the individuals to put effort on the task which generates their personal earnings and leads them to turn to the socially beneficial task. This would offset some of the adverse effects that the external incentives create on the effort towards socially beneficial task.

Teachers in a school are free to communicate among themselves about their past behaviors and the wage they have earned in recent past. This might induce them to change their behavior and follow what the other teachers have done in order to conform to a social norm. Past experimental evidence suggests that information about past behaviors of others leads individual choices to stick to a social or group norm. In our setting we test whether a treatment of information feedback about others' past behaviors induce them to conform to the group behavior and whether this leads to a shift in the effort allocations between the two tasks in any significant way.

Our findings show that incentivizing personal earnings substitutes away efforts from the socially beneficial task towards the task that generates more to their personal earnings as is predicted by the standard theory on multi-tasking principal-agent models. Apart from this we do not find any additional effect of cross-crowding out of intrinsic motivation where external incentives in the personal task would crowd out intrinsic motivation to invest in the socially beneficial task. This effect has been claimed by some of the teachers' unions, but our data reveals no such additional effect of crowding out on investment choices. In the context of teachers our result shows that the claims of the teachers' unions cannot be justified though the findings conform to the potential issues raised by the studies in the education literature about the adverse effects that performance-pay can create on the overall human capital formation of the society. Our results suggest that the performance-pay policy might have an effect on teachers to skewing their efforts towards teaching to the test substituting other aspects of teaching like generating curiosity and other higher-order learning skills among students.

We also find that our treatment of non-deterministic earnings does not weaken the power of incentives in any significant way though it shows some evidence of a dynamic discouragement effect of exerting effort on the task that generates personal earnings. Information feedback on past behaviors of others leads individual choices to cluster around some group average. However, there is no effect in aggregate on individual effort allocations.

Our results suggest that under performance-pay policy based on students' test score there might exist certain potential problems of affecting overall learning and higher-order skill formation among students even if their test scores get inflated, a potential issue which the policymakers should be considering while designing the wage structure for teachers. The effects of random factors in the generation of test scores as well as the information about behaviors of other teachers cannot help to decrease the adverse effects that performance-pay can create. These results reconcile the claims that some of the studies in the education literature has been making. Apart from this our findings cannot justify the claims of the teachers' unions as it shows no evidence that a crowding-out of teachers' motivation can arise from high incentives on test scores.

The paper is organized in the following way. In section 2 we discuss the literature to which our result connects. In section 3 we set up a basic theoretical model to show the effect of incentives on an agent's effort allocation. In section 4 we discuss our experimental design and in section 5 we deal with the major hypotheses that we want to test. Section 6 reports our results and Section 7 concludes.

2 Literature

Our paper can be broadly related to four different strands of literature: 1) the empirical literature on the effect of merit pay on students' achievements, 2) experimental studies on multi-tasking principal-agent problems, 3) theoretical and experimental literature on motivation crowding-out, and 4) studies related to the effect of group norm and social identification. In what follows we review these four volumes of work separately.

A. Empirical Literature on Effect of Merit Pay

The first strand of literature is basically the different empirical findings on the effect of the merit pay system on students' standardized test scores. The basic question that all of this literature tries to answer is the effectiveness of the new wage policy on test score gains, which is different from the question our paper raises. We broadly study if the merit pay causes efforts on teaching to the test to go up. Even if there is an increase in test scores, is the overall education level of the society improved? So while the existing literature studies test-effectiveness of the policy, we study its effect on effort allocation. There is a mixed finding in the literature in terms of the the effects on test scores. Although some part of the literature indicates a positive impact of the merit pay system, but there are a few of them which show that the incentive system might not have significant impact on the students' scores.

Glewwe, Ilias and Kremer (2010) conducted a study at different schools in Kenya where the scores increased from the second year of the introduction of the merit-pay program, if not from the first year but the effect did not persist after the completion of the program. The evidence supports the hypothesis that pay-for-performance increase test scores in the short-run but have little pact on long-term learning of the students. This gives an indication of the multidimensionality of teachers' efforts and some crowding out of unobservable characteristics by teachers. Glewwe, et al. (2002a) evaluate a program that provided prizes to teachers in schools that preformed well on exams and had low dropout rates. Teachers responded to the program not by increasing attendance, but by increasing prep sessions designed to prepare students for the exams. The result is consistent with the model in which teachers responded to the program primarily by increasing effort devoted to manipulating test scores, rather than by increasing effort at stimulating long-term learning, test scores for pupils who had been part of the program initially increased but then fell back to levels similar to the comparison group at the end of the program. Atkinson et al.(2009) conducted a study in England where there was no clear evidence of improvement of student scores from introducing mert-pay for the teachers. Goodman and Turner (2010) sampled some 'high need' schools in the New York City under the program of incentive pay and the result again does not show evidence that the program improved student achievement. Although these papers give the hint that incentives are not a long term solution for student learning, but we cannot specifically learn anything of the relative allocation of teachers' efforts on different dimensions of teaching. For this reason we needed to conduct a controlled experiment and show the exact nature of the relative importance the teachers give in different dimensions rather than measuring through the imperfect measure of students' test scores.

On the other hand, Winters et al.(2008) reports a study conducted in Arkansas where the individual teachers were awarded bonuses based on their students' improvement on the Iowa Test of Basic Skills. The result showed that there was statistically significant improvement in all three subjects (math, reading, language) tested. Hudson (2010) conducted a study where introduction of bonus raised math scores relative to the control group by .15 SDs. A randomized evaluation study was conducted in India by Muralidharan and Sundararaman (2010) where after 2 years of introduction of the bonus program, students in incentive schools scored better than the control group by .28 SDs in math, and .17 SDs in language. Another study has been conducted by Lavy (2009) in Israel where there was a rank order tournament among teachers of each subject, with fixed rewards of several levels. Teachers were ranked based on how many students passed the matriculation exam, as well as the average scores of their students. There were overall improvements in pass rates in Math and English due to an overall change in teaching methods, increased after school teaching, and increased responsiveness to student needs among teachers. Another very recent paper by Woessmann (2014) gives cross-country evidence about the long-run effects of pay-for-performance on teacher sorting. They show that the general-equilibrium effects of performance-related teacher pay include long-term incentive and increase math, science, and reading achievement across countries. Scores in countries with performance-related pay are about one quarter standard deviations higher.

Our paper differs from this empirical literature in the sense that while they study the effects of performance-pay on the test scores of students, our study sheds light on the impact of the wage policy on teachers' effort allocation. We also test whether this incentive can create any crowding-out of intrinsic motivation for teachers.

B. Experiments on Multi-tasking Problems

The second strand of literature which our paper can contribute to is the experimental literature on the multi-tasking moral hazard problems. There are only a few experiments conducted in this issue, mostly devoted to static effects of bonus wage on effort substitution. Fehr and Schmidt (2004) study a problem where one task is contractible and they focus on the pros and cons of piece-rate versus bonus contracts. But in their paper the agent puts more effort in the piece-rate treatment than the bonus treatment as a result of a reciprocity towards the principal. But our paper differs from this in the sense that the agent in our paper gives more effort on the socially beneficial task out of his general altruism but not out of any reciprocal behavior towards the principal.

Bruggen and Moers(2007) show that active trade-off between effort level and effort allocation exists. They consider two tasks where effort in one task is easily measurable from the outcome while the effort in the other task is not.They study the effects of financial incentives on distortions in effort allocation. Hoppe and Kusterer(2010) compares the behavior of the agents when one single agent is assigned to do both the conflicting jobs to that when there are two agents each assigned with a single task. Our paper differs from the latter in that there is a single agent (quite similar to the basic framework of Holmstrom and Milgrom(1991)). Past experiments have only considered static effects of incentives on effort allocation while the effects of dynamics have been ignored. Our paper considers dynamics of individual behavior over and above their static behavior under risky situations as well as under effects of social considerations.

C. Motivation Crowding Out

The third strand of literature that we focus on is that of the motivation crowding out theory. There is an existing literature in psychology and economics that talks about extrinsic vs intrinsic motivation and supports the idea that providing incentives on a task may crowd out effort on that task as extrinsic incentives like monetary incentives can undermine intrinsic motivation. Although the existing literature deals with motivation crowding-out for a single task, our paper looks at this issue differently to study motivation crowding out in a multi-task framework. We do not find evidence of a cross-crowding out of motivation in our multi-task setting where external incentives in one task crowds out motivation for a different task.

In the psychology literature, Harvey(2005) shows that motivation crowding out occurs when explicit rewards are perceived as controlling, which results in individuals having greater satisfaction by not being intrinsically motivated. There are several other studies examining this issue, for example Ryan and Deci(2000), Cameron *et al*(2005), Vansteenkiste *et al*(2006) etc. For an extensive list of psychological studies on motivation crowding out, see Kunz and Pfaff(2002).

There is a vast literature in economics as well on the concept of motivation crowding-out. For example, Andreoni(1993) studies the effects of government tax on voluntary public goods provision. Although the questions are little different, but they all pertain to the basic theory that any external intervention might crowd out intrinsic motivations. Chang and Li(1999) and Benabou and Tirole(2000) derive theoretical foundations of the motivation crowding out in agency theory. In their model, the principal has superior information about the agent's ability to successfully accomplish a task. In such a setting, increased monetary incentives may backfire when interpreted by the agent as a signal of low competence, diminishing his self-confidence and his overall motivation to work vigorously. This is a phenomenon where the agents' preferences change due to external intervention by lowering their self-esteem and thus decreases their incentives to put more effort. This was called the 'hidden cost of reward' by Frey(1997). This was established further by a number of lab and field experimental studies (Gneezy and Rustichini, 2000, Gneezy and Rustichini, 1998, Zanella,1998) . Gneezy and Rustichini(2000) conducted a field experiment where under certain scenarios participants who were offered monetary incentives performed more poorly than those working for free. The basic idea is that the labor supply curve shifts to the left due the decrease in the intrinsic motivation factor. Gneezy and Rustichini (1998) also finds evidence of the crowding out theory but the theory works only when a substantial amount of money was provided as incentives. Zanella(1998) shows the possibility of a crowding-out effect for intrinsic motivation in the form of a tendency for reciprocal behavior. In summary, all of the literature focuses on crowding-out in case of a single task. Our paper looks into it in a different way. In place of a single task we examine the effects of crowding-out with two conflicting tasks.

D. Group Norm

There are some studies in the literature on the effects of group norm on individual behaviors. Kandori(1992) provides a theoretical model on the effects of social norm in enforcing cooperation among individuals in a repeated game. Akerlof(1980) derives a model to determine how individuals choose which norms to follow that do not contradict their economic gains. Fehr and Fischbacher(2004) studies different enforcement mechanisms for sustaining social norms. Andreoni and Bernheim (2009) shows that information and observability of past choices leads to a more pro-social behavior. An increasing number of studies use the concept of social norms to explain important phenomena (Lindbeck, Nybergand Weibull,1999, Solow, 1990). Experiments conducted by Mago, Savikhin and Sheremeta (2012) show adherence to the group norm by individuals when they receive information feedback on the behaviors of the group. Our paper relates to the latter study in the sense that we also test the effects of information feedback on individual effort choices. But while Mago *et al* (2012) studies effort choices for a single task, we show the effects of information feedback when individuals can substitute efforts in two different tasks.

3 Model

This section deals with theoretical models to address the effects of performance-pay on teachers' effort allocations in a multi-tasking principal-agent framework. We provide theoretical predictions and comparative-static analyses of an increase in performance-pay on optimal effort choices of the agents.The ways in which the earnings are generated for the agents are different in different models. Section 3.1 deals with the case when the earnings function is deterministic, while 3.2 deals with the case when there is a risk involved in generating earnings from one of the tasks. We show the cases where the bonus decreases efforts on teaching higher-order skills and the cases where it does not. We need these theoretical predictions so that we can compare the result with our experimental data. Section 3.3 deals with some issues on behavioral dynamics of agents' choices. Here we discuss how a discouragement effect coming from risk as well as conformity to a group norm affect their choices over time. In Section 6 we test experimentally whether these effects alleviate or aggravate the negative impacts of bonus on higher-order learning skills.

3.1 Deterministic Outcome

Our model relates closely to the multi-tasking principal-agent problem, introduced by Holmstrom and Milgrom (1991). In our case the principal is the government who cares for the overall education and human capital formation of the society while the agents are the teachers who receive wages from the principal. We consider multidimensionality of teachers' tasks inside a classroom(Hannaway,1991). In this model we assume that they have two tasks to perform, instead of a single task. They can engage in the task of generating basic test taking skills, or they may engage in the task of helping the students to promote creative and higher order thinking of the students. This distinction of different teaching roles that the teachers play is the notion of the 'multi-task' in Holmstrom-Milgrom (1991). Let the corresponding effort levels for the two tasks be e_1 and e_2 respectively. So the choice variable for the agents is not a single effort level, but instead a vector of effort levels $e = (e_1, e_2)$.

Let $0 \leq \underline{e}, \overline{e}, \leq 1$ denote the lower and upper limits of the effort levels that the agents can choose. Let the cost function of the agent for the two effort levels be

$$C = C(e_1, e_2)$$

with the assumptions $C_i \geq 0, C_{ij} \geq 0, i, j = 1, 2, i \neq j$. C_{ij} measures the extent to which the marginal cost of one effort increases with the increase of the other, i.e. in other words it measures how the two efforts are substitutable by the agents. The non-negativity condition on C_{ij} assures that the marginal cost can never be negative in our model. Relaxing this condition would result in the strategic complementarity case of the agents' cost function in the Holmstrom and Milgrom paper. But for our model we stick to the non-negativity restriction on C_{ij}. in our present scenario. If we think of e_1 and e_2 to be the time allocated by a teacher to the tasks 1 and 2, rather than efforts, this would seem to be a more realistic assumption since the total time a teacher gets for teaching complete lesson is fixed. For simplicity we assume that $C_{ij} = C_{ji}$.

We introduce in this model an additional feature that the agents also have an inherent motivation to perform the tasks. Let $\alpha(e_1, e_2)$ be the inherent utility that the agent derives from performing the tasks. α_2 can be interpreted as the marginal utility the teachers derive in generating higher-order thinking abilities of the students, helping them to learn the subjects thoroughly, and thus raising the human capital level of the society. This can also be their sole passion for teaching the subjects they are interested in. α_1, corresponding to the task of teaching to the test, may, on the other hand, be the marginal utilities the teachers derive when their own students perform well in the tests. So we can assume that $\alpha_i \geq 0, \alpha_{ij} \leq 0, i, j = 1, 2, i \neq j$.

Initially, when there is a fixed wage, the overall utility for the agents is ,

$$U(e_1, e_2) = s + \alpha(e_1, e_2) - C(e_1, e_2),$$

where s is the fixed salary paid to the agents. When the principal can observe the effort choices perfectly, we can reach the first best solution by constructing an incentive contract to induce socially ideal effort. We can form the standard concave welfare function of the society. The social welfare function is the overall learning and education of the society as a function of teachers' effort levels on the two tasks. This welfare function is also the principal's objective function:

$$W(e_1, e_2) = f(e_1, e_2)$$

with $f_i \geq 0$, and $f_{ii} < 0$, $i = 1, 2$. f_1 and f_2 are the marginal contributions of the two different tasks . We can assume that $f_1 < f_2$, *i.e.* the marginal contribution of e_2 is strictly higher than that of e_1.This means that when the students develop higher order thinking, their human capital is increased and the society benefits more than if the students learn only to perform well in the tests. These assumptions imply that the welfare function follows diminishing marginal welfare for each of the tasks and this guarantees a maximum.

The first order conditions to maximize the welfare function give the first best solutions of the effort levels. Since the marginal contribution of task 2 is greater than that of task 1, in the optimum the society should choose more effort in task 2 than in task 1.

In the case where the efforts are not observable, incentive problems arise. Let the principal offer a bonus rate b to incentivize effort from the agents. Let us suppose that the efforts are observable imperfectly through a function P , where P denotes the scores of the students in the test and it depends on both types of efforts.

$$P = P(e_1, e_2).$$

We let $P_1 > P_2$, *i.e.*, the marginal contribution of the teaching to the test effort on the scores of the students is higher than the marginal contribution of the other effort. Since the principal cannot observe the effort levels individually, he has to base his incentives only on the value of P, a function of both effort levels. Herein lies the moral hazard problem, Since $P_1 > P_2$, and the incentive structure is based on P, the agents are more willing to provide e_1 than e_2, but on the other hand the marginal contribution of e_2 on the social welfare is higher than that of e_1, which is given by the condition $f_1 < f_2$.

The agents' problem is to maximize :

$$Max_{e_1,e_2} s + bP(e_1, e_2) + \alpha(e_1, e_2) - C(e_1, e_2)$$

where s is the fixed wage independent of the effort choices. The principal has to set the bonus wage in such a way that the agent chooses the effort levels that are optimal for the principal. Thus the optimal effort levels that the agent chooses to maximize his own utility serves as the incentive compatibility constraint of the principal.

The participation constraint (PC) for the agent is

$$s + bP(e_1, e_2) + \alpha(e_1, e_2) - C(e_1, e_2) \geq \underline{u}$$

where $\underline{u}$ is some reservation utility level. The participation constraint ensures that the agent's utility has to be greater than some minimum level of a reservation utility. Along with this, there is also another constraint : $s + bP(e_1, e_2) \geq s$. This constraint is called the limited liability constraint (LLC) as it ensures that the bonus cannot be negative. We can assume that the value of $\underline{u}$ is so low that the participation constraint does not bind. So we can ignore the participation constraint, and the relevant binding constraints are the ICC and the LLC. If the LLC does not bind, the moral hazard problem would not have arisen, since the principal could have given any penalty to the agent in order to achieve the first best outcome.

The principal's optimization problem is thus :

$$Max_b f(e_1, e_2) - s - b(g_1 e_1 + g_2 e_2)$$

s.t. 1) *Incentive compatibility constraint* : $(e_1^*, e_2^*) = \arg\max s + bP(e_1, e_2) + \alpha(e_1, e_2) - C(e_1, e_2)$.

2) *Limited liability constraint* : $s + bP(e_1, e_2) \geq s$.

Solution to the problem is straightforward. But rather than the exact value of optimal bonus level, we are more interested in the agent's problem to see the comparative static analysis of the change in the two effort levels, with the change in b.

The first order condition of the agent's problem (or the principal's incentive compatibility constraint) gives:

$$bP_1(e_1^*, e_2^*) + \alpha_1(e_1^*, e_2^*) - C_1(e_1^*, e_2^*) = 0$$

and

$$bP_2(e_1^*, e_2^*) + \alpha_2(e_1^*, e_2^*) - C_2(e_1^*, e_2^*) = 0.$$

Assuming interior solutions the first order conditions give the optimum values of e_1 and e_2, i.e. e_1^* and e_2^* .

The first order conditions give the following comparative statics:

$$\frac{\partial e_1^*}{\partial b} = \frac{\begin{vmatrix} -P_1 & bP_{12} + \alpha_{12} - C_{12} \\ -P_2 & bP_{22} + \alpha_{22} - C_{22} \end{vmatrix}}{\begin{vmatrix} H \end{vmatrix}}$$

and

$$\frac{\partial e_2^*}{\partial b} = \frac{\begin{vmatrix} bP_{11} + \alpha_{11} - C_{11} & -P_1 \\ bP_{21} + \alpha_{21} - C_{12} & -P_2 \end{vmatrix}}{\begin{vmatrix} H \end{vmatrix}}$$

where $|H|$ is the Bordered Hessian which is strictly positive from the second order condition. Suppose we assume linearity of the P function to give a clear insight. This means that the second order derivatives and the second order cross partial derivatives of P are zero. From the comparative static exercise, we can see that if $\frac{P_1}{P_2} > \max\{\frac{\alpha_{12} - C_{12}}{\alpha_{22} - C_{22}}, \frac{\alpha_{11} - C_{11}}{\alpha_{21} - C_{12}}\}$, $\frac{\partial e_1^*}{\partial b} > 0$ and $\frac{\partial e_2^*}{\partial b} < 0$ i.e. e_1 rises and e_2 falls as a result of the introduction of the bonus. Since the second order condition ensures that $\frac{\alpha_{11} - C_{11}}{\alpha_{21} - C_{12}} > \frac{\alpha_{12} - C_{12}}{\alpha_{22} - C_{22}}$, the parameter range for this incentive problem to take place is

$$\frac{P_1}{P_2} > \frac{\alpha_{11} - C_{11}}{\alpha_{21} - C_{12}}.$$

The condition implies that if the marginal contribution of the effort on teaching to the test on students' standardized scores is very high compared to that of the effort on the other task, the teachers would substitute more efforts on teaching to the tests than generating other higher dimensions of skills. The ratio of the two marginal contributions has to be greater than a certain value for the moral hazard problem to take place.

It can be noted that there is some parameter range where the bonus might even be beneficial to the society. If $\frac{\alpha_{12}-C_{12}}{\alpha_{22}-C_{22}} < \frac{P_1}{P_2} < \frac{\alpha_{11}-C_{11}}{\alpha_{21}-C_{12}}$, increase in the bonus rate increases both the effort levels.

Under this parameter range an incentive structure is unambiguously better for the society. This is true keeping the motivation parameters $\alpha_i' s$ remaining the same. But if there is a crowding out of intrinsic motivation of the teachers, as is claimed by the teachers' unions, then α_2 decreases as a result of external incentives. This is a cross-crowding out of intrinsic motivation which results in a decrease in e_2 as a result of incentivizing e_1. If this crowding-out effect dominates effort increasing bonus effect, we can observe a net decrease of effort in task 2 for incentivizing task 1, i.e. $\frac{\partial e_2}{\partial b} < 0$. This is basically a leftward shift of the effort supply curve of the teachers due to decrease in α_2. Firstly, this effect is different from the negative direct effect of bonus under other parameter ranges. This indirect effect of bonus through the motivation parameter is valid for the other parameter ranges as well thereby only multiplying the negative effects of bonus. Secondly, this issue of cross-crowding out of motivation has not been addressed in the earlier motivation crowding out literature which only deals with crowding out for a single task. Some studies in the education literature which opposes the performance-pay for teachers has also not raised this issue. They only claim the negative substitution effect of bonus. So in our experiment we want to test if there is any evidence of this cross-crowding out because in that case the negative impact of bonus wage would be much more than that claimed by the education literature.

When $\frac{P_1}{P_2} > \frac{\alpha_{11}-C_{11}}{\alpha_{21}-C_{12}}$ the multi-tasking incentive problem for the teachers arises even in the absence of any motivation crowding-out. Under this parameter range the marginal contribution of teaching to the test effort on the student scores is very high compared to the other effort. The teachers have more incentives to give effort on teaching to the test, rather than helping the students to actually learn.

The proponents of the incentive scheme can argue that it is only under this parameter range that the incentive problems for the teachers arise. In the range $\frac{\alpha_{12}-C_{12}}{\alpha_{22}-C_{22}} < \frac{P_1}{P_2} < \frac{\alpha_{11}-C_{11}}{\alpha_{21}-C_{12}}$ both the effort levels increase with the increase in the bonus rate. When $\frac{P_1}{P_2} < \frac{\alpha_{11}-C_{11}}{\alpha_{21}-C_{12}}$ although e_1 decreases, but at the same time e_2 increases, and it is this effort level e_2 that the principal values the most. Thus for a wide range of parameter values the incentive scheme works well, and it is debatable that in the society the actual value of the parameter should happen to be always in the range $\frac{P_1}{P_2} > \frac{\alpha_{11}-C_{11}}{\alpha_{21}-C_{12}}$.

On the other hand, the opponents of the incentive scheme can justify that the above argument is true only for a fixed level of C_{12}. As the value of C_{12} increases the range $[\frac{C_{12}}{C_{22}}, \frac{C_{11}}{C_{12}}]$ would shrink more. At the extreme when $C_{12} = \sqrt{C_{11}C_{22}}$, $\frac{C_{12}}{C_{22}} = \frac{C_{11}}{C_{12}}$. This is a point of a knife-edge situation where e_1 and e_2 are perfectly substitutable. A marginal change of P_1 on either direction would imply the teachers will give full effort only in one task, and zero effort on the other. Obviously, this extreme situation is unlikely in the real context, but at least for a fairly high C_{12}, $\frac{C_{11}}{C_{12}}$ would very low and so it becomes more likely that $\frac{P_1}{P_2}$ exceeds the value and fixed pay is then better than bonus.

Let us now demonstrate with an example our idea to disentangle the effects of crowding-out of intrinsic motivation and substitution effect. Suppose the motivation function is a linear function of e_1 and e_2, i.e. $\alpha(e_1, e_2) = \alpha(k_1 e_1 + k_2 e_2)$, where k_1,k_2,$\alpha > 0$ are three constants. Suppose the wage function is also a linear function $W = s + be_1 + e_2$.Then the optimal values of e_1 and e_2 would be represented as

$$e_1^* = e_1^*(b, \alpha)$$

and

$$e_2^* = e_2^*(b, \alpha).$$

where under certain conditions $\frac{\partial e_2^*}{\partial b} < 0$ and $\frac{\partial e_2^*}{\partial \alpha} > 0$. With an increase in b, e_2^* decreases for any level of motivation parameter α. e_2^* can also decrease with decrease in the motivation parameter α. Thus with an increase in b we can predict the decrease in e_2^* with α remaining the same. From our experimental data we can get the actual decrease in e_2^* when there is a strict increase in incentive in task 1. If the actual decrease is greater than the decrease predicted by the change in b, then the residual decrease in e_2^* is due to the decrease in α, i.e. due to crowding-out of intrinsic motivation. But if the actual decrease is equal or less than the predicted decrease, then the change in b can explain the entire decrease in e_2^* and we would find no evidence of crowding-out of intrinsic motivation.

3.2 Risky Outcome

In this section we introduce risk into the outcome function.The outcome function is not deterministic in the sense that there is a random component involved between the effort choices and the students' score levels. Effort choices do not directly determine student score levels but score levels do increase on average with teacher efforts. This risk arises because the ability of the students may be unknown, or it may be due to some random events occurring during the examination time etc. We assume that the teachers have a constant absolute risk aversion (CARA) risk preference represented by the following negative exponential utility function,

$$U(e_1, e_2) = -e^{-\eta[W - C(e_1, e_2)]}$$

where W is the total wage that the agents or the teachers receive, and $C(e_1, e_2)$ is the total cost of providing efforts. $\eta > 0$ is the coefficient of absolute risk aversion($\eta = -\frac{U''}{U'}$). We assume for simplicity that the cost function is convex, i.e.

$$C(e_1, e_2) = \frac{1}{2}(c_1 e_1^2 + c_2 e_2^2) + \gamma e_1 e_2$$

where c_1, c_2 and$\gamma \geq 0$. W is the total wage which is determined by a linear score function of the students,

$$P = g_1 e_1 + g_2 e_2$$

with $g_1, g_2 \geq 0$. We assume here that this outcome function is not deterministic as the $g_i's$ are unknown. The agents do not know for certain about the exact mapping from their effort choices to the outcome. We assume that $g_i\text{~}N(\mu_i, \sigma_i^2)$, $i = 1, 2$. The wage function is the linear function of the score levels

$$W = s + b(g_1 e_1 + g_2 e_2).$$

So the von Neumann-Morgenstern expected utility function of the agents become

$$E(U) = E(-e^{-\eta[s+b(g_1e_1+g_2e_2)-\frac{1}{2}(c_1e_1^2+c_2e_2^2)-\gamma e_1e_2]})$$

Note that we have assumed away the intrinsic motivations of the agents in this subsection. This is only to simplify the calculations. The motivation parameters $\theta_i' s$ are modelled such that they would not interact with the risk parameters, and so this would not change the results qualitatively.

We know that if x follows $N(\mu, \sigma^2)$, then $E(e^{tx}) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$. Therefore, the expected utility of the agents is equal to the value, multiplied by -1, of the moment-generating function of a normal distribution with mean $b(\mu_1 e_1 + \mu_2 e_2)$ and variance $\frac{1}{2}\eta^2 b^2(\sigma_1^2 e_1^2 + \sigma_2^2 e_2^2)$. The expression for the expected utility function for the agents is

$$E(U) = E(-e^{-\eta[s-\frac{1}{2}(c_1e_1^2+c_2e_2^2)-\gamma e_1e_2+b(\mu_1e_1+\mu_2e_2)-\frac{1}{2}\eta b^2(\sigma_1^2e_1^2+\sigma_2^2e_2^2)]})$$

Applying a monotone transformation $U \to -\frac{1}{\eta}\ln(-U)$, we have the following final expression for the expected utility:

$$E(U) = s + b(\mu_1 e_1 + \mu_2 e_2) - \frac{1}{2}\eta b^2(\sigma_1^2 e_1^2 + \sigma_2^2 e_2^2) - \frac{1}{2}(c_1 e_1^2 + c_2 e_2^2) - \gamma e_1 e_2.$$

The agents' problem is thus to choose e_1 and e_2 to maximize the expected utility. We restrict ourselves here only to the interior solutions. The corner solutions can be easily derived similar to the previous subsection of the deterministic case. The first order conditions give the interior solutions as

$$e_i = \frac{b^3\mu_i\eta\sigma_j^2 + b\mu_i c_j - \gamma b\mu_j}{(\eta b^2\sigma_i^2 + c_i - \gamma^2)(\eta b^2\sigma_j^2 + c_j)}, \; i, j = 1, 2, \; i \neq j.$$

It can be verified that $\frac{\partial e_i}{\partial \sigma_i^2} < 0$, and $\frac{\partial e_i}{\partial \sigma_j^2} > 0$ iff $\frac{\mu_i}{\mu_j} > \frac{c_{ii}}{c_{ij}}$. Thus the complementarity of the two actions implies that as risk increases for a certain task, individuals decrease their allocation of effort from that task. Efforts on the other task increases if the average marginal contribution of that task is above a threshold level.

The relation between e_i and b is a little tricky. If $\mu_1 = g_1$ and $\mu_2 = g_2$, and $\sigma_1, \sigma_2 = 0$ in the expression of e_i, we get the exact condition of the deterministic case, i.e. e_2 falls with increase in the bonus rate if $g_1 \geq \frac{c_1 g_2}{\gamma}$. We introduce risk into the model with the $g_i's$ being drawn from normal distributions with means $\mu_i = g_i$, and variance σ_i^2. Thus the risk is introduced as a mean-preserving spread of the deterministic case, with $g_1 \geq \frac{c_1 g_2}{\gamma}$. In that case we require an additional restriction on σ_1 for e_2 to fall with increase in bonus rate, the restriction being that σ_1 has to be greater than a minimum threshold level of say $\overline{\sigma_1}$. Thus there is a trade-off between the two conditions, i.e. as we increase μ_1 above g_1, $\overline{\sigma_1}$ decreases, i.e. the condition on the variance becomes less and less restrictive.

3.3 Behavioral Dynamics

The theory predicts that an increase in bonus pay increases efforts allocated to the task in which there is a direct external incentive. The effect on the other task can come from either the direct substitution effect of bonus wage or indirectly through the crowding-out effect. The substitution effect can be positive or negative depending on the parameter range, while negative cross-crowding out of intrinsic motivation can decrease effort on this other task even in presence of positive substitution effect. Although the theoretical model presented above is a one-shot model, we can also assume an intertemporal behavior of the teachers from this model. Teachers give efforts to teach their students throughout an academic year, and at the end of the year a bonus wage accrues to their payment. Teachers are free to communicate among themselves about their salaries, and maybe also about how they taught in the classroom. The first broad question that we try to answer is whether the behavior of the teachers can change dynamically as they get to know of the past behaviors of their colleagues. Literature shows that people generally have a tendency to stick to a perceived social norm (Akerlof, 1980, Kandori, 1992 etc.). So does the information about group behaviors change individual perception about social norm (or group norm) and affect their current behavior? For example, does the information that my fellow colleagues have given more effort choices to teaching to the test, and thus have earned more money induce me to give more effort on teaching to the test, even if I, by myself, had initially preferred generating higher order thinking abilities of the students? This implies that for me there is a change in my view of the social norm. For example, suppose a teacher initially has a prior estimate of the social norm. So in the beginning a teacher may act according to this prior social norm. Subsequent information about others' behaviors might change his posterior estimate of the social norm and thus consequently his own behavior.

In quantitative terms, this can be explained through the cost function of an agent at time period t :

$$C(e_1, e_2) = \frac{1}{2}c_1 e_1^2 + |S_1^t - e_1^{t-1}|k_1 e_1 + \frac{1}{2}c_2 e_2^2 + |S_2^t - e_2^{t-1}|k_2 e_2 + \gamma e_1 e_2,$$

where S_1^t is the agent's estimate of the social norm task 1's effort choice at period t . The same is for S_2^t. We assume that the agent starts with prior estimates of the social norm, say S_1^0 and S_2^0 at the beginning of period 1. Subsequently his estimate of the social norm changes as he receives information about other agents' behavior. The posterior estimate of the social norm can be seen as a function of the average effort level of the other agents in the earlier period. Thus if an agent knows that everybody has given more effort in a certain task, he gets encouraged to exert more effort in that task from future periods as he wants to stick to the social norm.

There can be a number of possible measures of a social norm. One example can be the social or group average, i.e. $S_i^t = \sum_j e_i^{t-1}$, $i = 1, 2$, where j is the number of people in the cohort. In this case, if people try to stick to the group average, in aggregate the average behavior remains the same. But as the effort choices become more closer to the average, there is a decrease in variability of the choices. From teachers' perspective, we can say that this leads to a decrease in inequality as far as the level of teaching all the students receive from their teachers. Another example of social norm can be a minimum of the group choices, i.e. $S_i^t = \min e_i^{t-1}$, $i = 1, 2$. This leads to an decrease of everyone's choices to the minimum person's past choices if they perceive the minimum behavior to be the social norm. Likewise, the social norm can be a maximum of the group choices which would lead everyone to increase their efforts to the maximum person's past choices. In our experiment we want to test whether information about group behavior induce them in any way to stick to some social norm and if they do so what measure they use as social norm. We also test whether this conformance to social norm has any effect on the effort choices on the two tasks depending on their choice of the social norm.

The next environment we wish to consider is the effect of a dynamic discouragement in presence of risk in the outcome function. The theory gives a comparative static prediction that an increase in risk in any one of the tasks leads to a shift in the effort allocation to the other task. In our original motivating question we claim that the relation between teachers' effort choices and the outcome (i.e. the student scores) is not deterministic as there are some random noises that are involved. Apart from a comparative static effect of risk into the effort allocation, there might also be a dynamic effect in this situation arising from a discouragement effect of a risky environment. Consider from Section 3.2 the effort in task 1 i.e. e_1. For a high enough variance of outcome on that task, the actual realizations of g_1 can fall well below the mean of the distribution from time to time. Thus even a high effort choice on task 1 might not result in a satisfactory outcome. A repeated occurrences of this might lead to a discouragement among the agents to give effort on that task. This is the discouragement effect of risk on effort choices. Low realizations of the risky variable discourages them more than any encouragement they get from high realizations of the risky variable.

Like our previous case of teacher communication, we can present it through a modification of the cost function

$$C(e_1, e_2) = \frac{1}{2}c_1 e_1^2 + (\mu_1 - g_1^{t-1})k_1 e_1 + \frac{1}{2}c_2 e_2^2 + (\mu_2 - g_2^{t-1})k_2 e_2 + \gamma e_1 e_2,$$

where μ_i is the mean of the distribution of g_i and g_i^{t-1} is the realization of g_i in the previous period. As the realization of g_i falls below the mean, discouragement effect comes into play. The choice of mean of the distribution as the point of departure is not fully arbitrary for a symmetric distribution.

It is worth noting that this effect is different from the one-shot comparative static effect of risk predicted by the theory. Even in the absence of any discouragement, risk (measured by the variance) can itself cause the agents to provide lesser effort. But we can predict that apart from the pure risk effect, if there is an additional discouragement effect on the agents which accumulate over time and can thus reduce the incentive dynamically.

In the next sections we discuss in detail about the design of the experiment that we conduct, the hypotheses that we test and the corresponding results.

4 Experimental Design

In our experiment the subjects participated in a multi-tasking principal-agent game over a total of 25 periods divided into 3 stages. Earnings from the game were denominated in Experimental Currency Units (ECU), which were converted to money at a rate of $1 for every 100 ECU's.

In each round of a session the subjects had two different projects A and B in which they were allowed to invest. In each round they were able to invest up to 10 tokens into each project. Their investment decisions would affect their own earnings as well as have the possibility of generating earnings to some charity. At the first screen they were allowed to choose for which charity they wished to generate earnings. The three charities they were able to choose from were the local chapters of The American Red Cross, the American Society for Prevention of Cruelty to Animals and Habitat for Humanity. They could choose any of the three given charities or they could choose a 'None of These' option. If someone chose the latter, his actions did not generate any earnings to any charity. Those potential earnings were not paid to anyone. This was clearly explained to the subjects when they made this choice. The reason for allowing this 'None of These' option comes from a revealed preference argument that if a subject has chosen a certain charity with choosing none of them as a potential option, that means he actually cares for that charity to be given money. We read out to them the missions of each of the charities as reported by the charity review website. CharityNavigator.org. We also handed out to them printouts of a few pages of each organization's website that provide standard information regarding what donations are used for. This was done to inform them of what charities do to activate any good feelings towards charities so that they might gain utilities from their respective charities receiving money.

The design of the experiment is both within-subject and between subject. There are four different treatments. The treatments differ from each other in terms of how the personal earnings are generated. The following table shows the detailed design of the experiment and its different Stages.

	Stage 1 Periods $(1-5)$	**Stage** 2 Periods $(6-10)$	**Stage** 3 Periods $(11-25)$
Baseline	Equal Incentive	Bonus	Bonus
Risk	Equal Incentive	Bonus	Bonus + Risk
Information	Equal Incentive	Bonus	Bonus + Information

Table 1: Table showing Experimental Design

The first row describes the Baseline configuration where the marginal contributions to their personal earnings are almost similar for both the projects (denoted as Equal Incentives in the table) in Stage 1 and in Stages 2 and 3 a bonus incentive was given to the subjects for investing in project

A (denoted as Bonus in the table). The second row is the Risk configuration where Stages 1 and 2 are similar to that of the Baseline configuration, but in stage 3 there is an element of risk involved in earnings for project A. The third row is the Information configuration where Stages $1, 2$ and 3 have same incentives as the Baseline but in Stage 3 the subjects are informed regarding the investment choices of other subjects in their prespecified groups after each round. Each configuration with the corresponding earnings and charity functions are explained in detail as follows.

Baseline Configuration:

In Stage 1 which consisted of rounds 1-5, all subjects played the investment game where the benefit of investing in projects A and B is:

$$Benefit(a, b) = 55 + a + b,$$

i.e. the subjects get a base wage of 55 ECU's and 1 ECU for each unit investment in tasks A and B. The cost of investing in A and B is:

$$Cost(a, b) = 0.8 * a^2 + 0.6 * b^2 + 0.5 * a * b,$$

where a and b are the number of tokens invested in project A and B respectively. Their payoff in each round is the difference between the benefit and the cost. We explained to the subjects in detail about the cost function with special emphasis on the nature of the marginal cost on each project that depends on how many tokens one invests in the other project. They were also told at the beginning of the stage that this payoff function would vary from stage to stage. The predicted investments in A and B are 0 and 1 respectively.

Earnings generated for the charity are determined according to the following function:

$$Charity(a, b) = 2 * a + 8 * b.$$

Note that while there is an almost equal marginal gain in their personal earnings from investing in either project, investing a token in B would generate earnings to the charity 4 times more than investing a token in A. While they were making their choices, on the right hand side of their screen all the personal earnings and charity earnings calculations were auto-updated for them so that they did not have to worry about calculating the complex functions. For a clear understanding of the cost function the subjects were provided with a cost table where the cost calculations were done for the for each possible combination of investment choices. The subjects could refer to this table while making their decisions. At the end of every period in Stage 1, subjects could see their own investments in both the projects, their personal earnings as well as the earnings generated to the charity for the period. If a subject values the charity receiving money, he could choose an investment in B above 1 with how far depending on the intensity of his preference.

Stage 2 took place over periods $5-10$. The only difference with the earlier stage is in the marginal incentive on project A. There is now a monetary incentive to invest in project A, as the marginal benefit of investing in A is now 10 times that of investing in B. The cost function remains the same as in the previous stage. The payoff function is given as follows:

$$\pi(a, b) = 55 + 10 * a + b - 0.8 * a^2 - 0.6 * b^2 - 0.5 * a * b.$$

The earnings generated for the charity remains the same. So, a token invested in project A would increase their personal earnings more than what it would contribute to the charity. However investing in project B would give more ECU's to the charity than what it contributes to their personal benefit. After every period in Stage 2 subjects saw the same information as before about their own investments in both the projects, their personal earnings as well as the earnings generated to the charity for the period. The predicted values of investments in A and B are 6 and 0 respectively assuming no intrinsic motivation to invest in B. The number of tokens invested to project A should increase and that in project B should decrease compared to that in the previous Stage. This is because the external incentive on A would substitute their effort levels away from B. The payoff function was chosen in such a way that the fixed income of 55 ECU's remained the same as in the previous Stage. Additional to that an incentive of 10 ECU's to a token invested in A, the average earning of the subjects would be higher compared to the previous Stage. Clearly there is an income effect which acts in the direction of increasing tokens in either project. So if we get our predicted result of crowding out of project B in this Stage negating the income effect our result would be more robust.

Stage 3 lasted from period 11 through the final period 25. The subjects played the same game with the same payoff and charity functions as in Stage 2. This was set up so that the effects of different treatments can be compared with their behaviors in Stage 3 of this baseline treatment. In this treatment, at the end of Stage 2 we told the subjects to move on to Stage 3 where the set up is exactly the same as the one they played in the earlier stage. Subjects likely did not see a reason to refer to Stages 2 and 3 separately but we did so as to ensure a similar break between periods 10 and 11 in all treatments.

Risk Configuration:

In this treatment, the first two stages are exactly the same as the Baseline. The difference with the earlier treatment come in Stage 3 where we changed the payoff functions of the subjects in order to introduce risk into the setup. The payoff function of the subjects in this treatment is:

$$\pi(a, b) = 55 + 7 * a + x * a + b - 0.8 * a^2 - 0.6 * b^2 - 0.5 * a * b,$$

where x is the risky variable drawn randomly from a $Uniform[0, 6]$. If we compare this payoff function with that in Stage 2, we can see that if $x = 3$, we get back to the payoff function in Stage 2. Thus the risky payoff profile is a mean-preserving spread of the Bonus structure from Stage 2. We thus increase the variability of one portion of the earning from investment in A keeping the average to be the same. The subjects could not see the actual realizations of the variable while they made their investment choices. At the right hand side of the screen where their personal and charity earnings were auto-updated, they could only see the range of earnings that could be generated for themselves and for the charity while they made the choices. At the end of each period, the subjects could see the realizations of the risky variable in that round along with their investments in both the projects, their personal earnings as well as the earnings generated to the charity for the period. To further illustrate this, at the beginning of the Stage we took them through the software with a practice example. The exact payoff function in the practice example differed from the actual payoff function in this stage.

Information Configuration:

The last configuration is the Information configuration where the first two stages were exactly the same as in the two earlier configurations. The only difference with the others is that the subjects were randomly assigned into groups of five at the beginning of each session and after the end of each period in Stage 3 the subjects could see the choices of the other members in their own groups. This was told to them at the beginning of the Stage. The payoff functions that the subjects faced were also the same as in the Stage 2. The basic aim was to test whether the subjects had a tendency to follow the group norm and the group as a whole could coordinate to a single norm. We provided the subjects with a screenshot of a sample result screen where they see the table showing the information of the other group members. They were also told that choices that they could see in that sample table do not in any way give an indication of how they should play in the game. The numbers denoting the choices in both the projects in the sample table were randomly generated to avoid any bias.

We conducted 6 sessions of this experiment for a total of 75 subjects. All the sessions were conducted at the Laboratory for Research in Experimental Economics (LREE) at Southern Methodist University. Average total earnings were \$ 35 (including a \$5 participation fee) for an experiment lasting 1.5 to 2 hours. The software used for this experiment was programmed in Z-Tree, Fischbacher (2007).

5 Hypotheses

The theoretical model and the experimental design described above provides the support for several testable hypotheses. The first one stated below is concerned with the comparative static effects of incentives in the allocation of investments in the two projects.

Hypothesis 1: *Investment in project A increases and investment in project B decreases in Stage 2 compared to that in Stage 1 for all the configurations.*

The theoretical model in Section 2 shows that with introduction of bonus wage investments in A increase and investments in B decrease under certain parameter restrictions. However if the decrease in investment in B is greater than the theoretical predictions we can claim that external incentives in project A have a negative crowding out effect on intrinsic motivation of the subjects for investing in B.

The next broad question we want to answer is the effect of risk on the investment allocations. The introduction of risk leads us to two major effects on the investment choices, the effect of pure risk as well as the effect of a discouragement due to risk on the teachers' effort choices. This effect of pure risk (separated from discouragement effect) is again a comparative static effect on investment allocations. As risk increases in one task, the agents would shift the allocation of tokens away from that task to the less risky one. This is irrespective of whether they have any intrinsic preference. To elicit this effect, we compare the Stage 3 of the Baseline treatment treatment to that of the Risk treatment. This comparative static prediction is presented formally in the next hypothesis :

Hypothesis 2: *When we compare the data from Stage 3 of the Risk to that of the Baseline configuration, investment in project A decreases.*

Next we come to the test of the discouragement effect which is a dynamic effect. This is discussed in detail in Section 3.3. We conduct the Stage 3 of the Risk configuration for 15 periods, and try to study the dynamic effect of risk on the effort choices. After every period of this Stage, the subjects are informed about the realizations of the risky variable x of that period. We study the effect of discouragement on dynamic investment allocations of the agents. This can be stated formally in the hypothesis below :

Hypothesis 3: *During Stage 3 of the Risk configuration when the subjects know after every period the past realization of the risky outcome, the number of tokens invested into Project A decreases over time.*

The final hypothesis is regarding the case when the teachers have the full information of the past choices of their fellow colleagues. We compare Stage 3 of the Information configuration to that of the Baseline configuration to get the desired effect.

We conduct this Stage for 15 periods. After every period of the session, we provide each subject with the investment choices for each of the two tasks by the other subjects in the room. This leads us to our next hypothesis stated formally.

Hypothesis 4: *In the Information treatment, when the subjects are given the full information about the past behaviors of their group members, the investment choices within a group are more similar to each other compared to that in the Bonus treatment.*

We cannot compare Stage 2 and Stage 3 of both the Risk and the Information configuration while we are measuring the impact of either the risk or the information on the investment allocations. This is because, in Stage 2 the subjects would accumulate some wealth effects over the Stage. Thus when we transfer to Stage 3, this wealth effect from Stage 2 would translate into their behavior patterns in Stage 3 as well. So we cannot isolate this wealth effect from the behavior of the subjects in Stage 3. So, a direct comparison between Stage 2 and Stage 3 would not give a correct estimate of the effect of information as this wealth effect would either over or under estimate the total effect depending on which Project we are considering. So we need to compare Stage 3 of the Risk configuration to the Stage 3 of Information configuration in order to get the correct estimate.

6 Results

6.1 Overview

We begin by presenting the basic summary statistics of the performances of the subjects across treatments. Table 2 summarizes average investment levels in project A and project B for comparing two treatments, Equal incentives and Bonus, i.e. this is the comparison of the investment levels between Phase 1 and Phase 2 of all the sessions that we ran. According to standard theoretical prediction in Section 3, when subjects do not have any charity preference, investment in A increases and investment in B decreases with the introduction of the Bonus treatment, compared to the the Equal Incentives treatment. The subjects are incentivized to invest more into the project which has more marginal contribution towards their own earning rather than the charity earning.

	Treatments	
	Equal Incentives	**Bonus**
Tokens invested_A	1.84 (0)	6.03 (6)
Tokens invested_B	2.90 (1)	1.80 (0)
Personal Earning	49.30 (55.4)	80.04 (86.20)
Charity Earning	26.88 (8)	21.40 (12)

Table 2: Average tokens invested and ECUs earned
Predicted values are in parentheses

Table 2 shows that while the prediction of investment in B is 1 without any charity preference, our data shows that the subjects have invested much higher than that. This indicates that the subjects do have some preference for generating earnings to the charity. Compared to the Equal Incentives treatment, investment in A increases and investment in B decreases in the Bonus treatment. The decrease of investment in B from 2.9 to 1.8 can either be only due to the substitution effect of incentives, keeping their charity preference constant across the treatments, or it may be the case that some crowding-out effect is entangled into it. This necessitates an estimate of the effect of cross-crowding out of charity preference across the treatments.

Suppose the intrinsic motivation parameter in the utility functions of the subjects be α. The utility function of the subjects for Equal Incentives and Bonus treatments can then be written as:

$$\pi(a,b) = 55 + a * w + b + \alpha * (2a + 8b) - .8 * a^2 - .6 * b^2 - 0.5 * a * b$$

where w is the rate of external incentives for investment in B. The value of w is 1 in the Equal Incentives treatment and 10 in the Bonus treatment. Maximization of this payoff function gives the solutions for a and b as:

$$
\begin{aligned}
a &= 0.72w - 0.96\alpha - 0.3 \\
b &= 7.07\alpha - 0.3w + 0.96
\end{aligned}
$$

With an increase in w, investment in B decreases for any level of charity preference parameter α. Since the value of w increase by 9 ECU's across the two treatments (from 1 ECU in Equal incentives treatment to 10 ECU's in the Bonus treatment) the predicted decrease in investment in B is 2.7 tokens from the substitution effect with α remaining the same across the treatments. But the observed decrease in investment in B is 1.1 (2.9 in Equal Incentives to 1.8 in Bonus) which is less than the prediction. Note that investment in B can also decrease through a decrease in the charity preference parameter α. But as our data shows, the negative substitution effect from change in w can explain, on average, more than the observed decrease in investment in B regardless of the charity preference. So on average we do not see any additional crowding out effect on investment in B.

Since there may be some subjects who initially invest less than 2.7 tokens in the Equal Incentives treatment, for them the observed decrease in investment in B across the treatments may be less than 2.7 due to the lower bound of 0 on investment choices. That may drive down the average decrease in investment choices in project B. This leads to examining the decrease in investment for subjects whose initial investments are sufficiently high (i.e. for those subjects for whom there is no lower bound issue). Figure 1 shows the kernel density estimate of the distribution of difference in tokens invested in B across the two treatments for those subjects whose initial investment was greater than 4 tokens in the Equal Incentives treatment. As the figure shows, for approximately 60% of these subjects the difference in investment is 2-3 tokens, which conforms with the theoretical prediction. This shows that apart from the negative substitution effect of external incentives there is no additional crowding out effect even for those subjects whose initial investment is sufficiently high. We also run a t-test of whether or not the average shift in B is equal to 2.7 among that group against the alternative that it is greater than 2.7. We fail to reject the null hypothesis (mean= 2.9, p-value= 0.2).

Table 3 summarizes the same variables, investments in A and B, Personal Earning and Charity Earning comparing the treatments Bonus, Risk and Information. It compares the data from Phase 3 of all the sessions that we ran. The theory in Section 3 predicted that with introduction of risk into the outcome function investment in A increases and investment in B decreases compared to the Bonus treatment, under some parameter restriction. Also Personal Earnings decrease while and Charity Earnings increase in the Risk treatment compared to the Bonus treatment. But Table 3 suggests no treatment differences as the numbers remain mostly the same. So we need proper tests to have a definite conclusion. Again, with introduction of Information treatment we find there is no effect as such on average on investments in A and B.

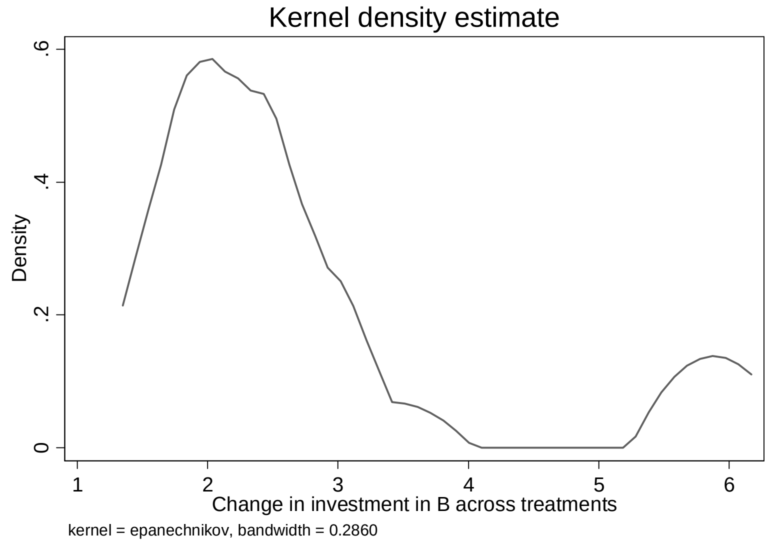

Figure 1: Kernel density estimate of the distribution of change in investment choices in B across treatments (Equal Incentives and Bonus) for individuals whose average investment is greater than 4 tokens in the Equal Incentives treatment

	Treatments		
	Bonus	**Bonus + Risk**	**Bonus + Information**
Tokens invested_A	6.13 (6)	5.98 (6)	6.01 (6)
Tokens invested_B	1.93 (0)	1.58 (0)	1.84 (0)
Personal Earning	80.04 (86.2)	77.15 (86.2)	81.50 (86.2)
Charity Earning	27.68 (12)	24.61(12)	22.50 (12)

Table 3: Average tokens invested and ECUs earned
Predicted values are in parentheses

6.2 Bonus

The introduction of the Bonus increases investment in project A and decreases investment in project B relative to the Equal Incentive treatment for all levels of intrinsic preference. Our data shows that the average number of tokens invested in project A of 6.03 in Bonus treatment is higher than that of 1.84 in Equal Incentives treatment, and the average tokens in B of 2.9 in Equal Incentives is higher than that of 1.8 in Bonus. Earnings to charity decreased from 26.5 to 21.4 with the introduction of the Bonus treatment. For our first set of basic statistical results we provide a set of non-parametric Mann-Whitney rank-sum tests on whether average investments differ between the treatments. For each test an observation consists of the average investment of all the individuals in each period for each treatment. So there are 5 observations in each of Equal Incentives and Bonus treatment. We find that there is a significant increase in investment in A and decreases in investment in B compared to the Equal Incentives treatment (both the p-values=0). The earnings to charity decreases in the bonus treatment (p value= 0.07). Given that we are aggregating across all periods, this test represents a very conservative test of whether or not there were differences across treatments.

For a more detailed analysis of treatment effects on investments in the two projects we run a regression with a treatment dummy which takes the value 1 if it is the Bonus treatment. Table 4 shows the result. In order to accommodate any subject level heterogeneity we use a fixed effect panel structure with the fixed effect on the individual level.

	InvestA	**InvestB**	**charity**
Bonus	6.199^{***}	-1.467^{***}	-10.25^{***}
	(0.286)	(0.336)	(2.670)
Period	-0.366^{***}	0.0674	2.783^{***}
	(0.0450)	(0.0528)	(0.420)
Constant	2.756^{***}	2.734^{***}	14.48^{***}
	(0.148)	(0.174)	(1.384)
Observations	748	748	748

Standard errors in parentheses

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 4: Regressions of investments in A and B and charity earnings from Bonus treatment

The results of the regressions show that with the effect of the Bonus treatment investment in A increases and investment in B decreases. Thus there is a negative effect on the investment in B due to the Bonus treatment. This is explained by the standard substitution effect of incentives. The overall charity contribution also decreases with the introduction of Bonus treatment. For all the three variables, the Bonus treatment dummy remains a significant predictor. Though effect on the increment in A(6.199) is much higher than that on the decrease of investment in B(-1.467), overall charity earnings decrease drastically with the Bonus treatment.

	InvestA	**InvestB**	**charity**
Bonus	5.232^{***}	-2.841^{***}	-10.32^{***}
	(0.141)	(0.325)	(2.540)
Period	-0.381^{***}	0.0693	2.791^{***}
	(0.0446)	(0.0513)	(0.410)
Constant	2.841^{***}	2.791^{***}	14.43^{***}
	(0.146)	(0.164)	(1.377)
Observations	187	187	187

Standard errors in parentheses
$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 5: Regressions of investments in A and B and charity earnings from Bonus treatment for individuals who invest greater than 4 tokens in Equal Incentives treatment

In Table 5 we run a separate regression similar to Table 4 but only for those individuals who have invested more than 4 tokens in the Equal Incentives treatment. This to test whether there is any cross-task crowding out effect of Bonus treatment on investment in B. The theory predicts that due to the substitution effect, the average decrease in investment B from Bonus treatment is 2.7 tokens. We want to test whether for these individuals the decrease in investment is significantly different from 2.7 tokens. There may be some subjects who initially invest less than 2.7 tokens in the Equal Incentives treatment. For them the observed decrease in investment in B across the treatments may be less than 2.7 due to the lower bound of 0 on investment choices. So we include only those subjects who have invested sufficiently high in the Equal Incentives treatment, and thus do not have the lower bound issue. Table 5 shows that the coefficient of Bonus treatment for the regression on investment in B is -2.841. A post estimation test shows that the coefficient is not statistically significant from 2.7 (p-value=1). Thus the decrease in investment in B is not significantly different from the theoretical prediction of substitution effect. This shows that there is no evidence of an additional cross-task crowding out effect due to the Bonus treatment.

6.3 Risk

Hypotheses 2 predicts that the introduction of Risk treatment should decrease investment in project A and increase investment in project B when we compare with Stage 3 of the Baseline treatment. Examining the impact of risk on investment allocations we find that data are not consistent with this hypothesis. Table 3 shows that there is no significant evidence in favor of a decrease in the average investment in project A and a increase in the average investment in B from Bonus treatment to the Bonus+Risk treatment. A non-parametric Mann-Whitney test also finds the differences across treatments not to be statistically significant(p-values $1, 1$ and 0.14).

	InvestA	**InvestB**	**charity**
Risk	-0.145	0.479^{**}	4.121^{***}
	(0.134)	(0.201)	(1.560)
Period_high	-0.361^{**}	-0.475**	-3.078^{*}
	(0.156)	(0.234)	(1.818)
Risk*Period_high	-0.295	0.347	2.186
	(0.198)	(0.296)	(2.303)
Constant	6.055^{***}	2.11^{***}	29.01^{***}
	(0.103)	(0.154)	(1.201)
Observations	795	795	795

Standard errors in parentheses
$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 6: Regressions of investments in A and B and charity contributions from Risk+Bonus treatment

For more detailed analysis we run separate regressions with treatment dummy controlling for individual level fixed effects in order to separate out any individual level heterogeneity. We create a dummy variable Period_high to denote periods from 20-25. Table 5 shows that the relationship between Risk + Bonus treatment and investment in the risky project A is not statistically significant .This along with the fact that the Period_high dummy has a negative significant coefficient gives an idea that instead of a static negative effect of risk on investment in A, there is some sort of dynamics present in the choices of the individuals. This dynamic behavior will be captured in the regressions in Table 6. The charity earning has a negative relation with the treatment effect. As the subjects decrease investments in both the projects with introduction of risk in project A, the effect on overall earnings to charity is negative.

In order to capture a discouragement effect we run a set of regressions for the effect of past realizations of the unknown variable in the Risk treatment on investments and charity earnings. We divide the range of values of the unknown variable $(0-6)$ into high-valued realizations if they take values 4, 5,6, and low-valued realizations if they take values $0, 1, 2$. We want to examine what effects these high and low realizations in the previous period have on investments and charity earnings. We create two dummy variables indicating high(Highval_{t-1}) and low realizations(Lowval_{t-1}) respectively. There might also be the case that the result is confounded by the wealth effect of the subjects. Higher income might induce the subjects to invest more tokens into either project. To disentangle the wealth effect we create two dummy variables which are the interactions of the Profit variable with the high and low value dummies respectively. Table 6 shows the effect of these high and low past realizations on investment allocations.

	InvestA	**InvestB**	**charity**
Highval_{t-1}	0.267	-0.441^{**}	-2.896^{**}
	(0.206)	(0.216)	(1.452)
Lowval_{t-1}	-0.469^{***}	0.191	0.125
	(0.163)	(0.227)	(1.527)
Period_high	-0.0224	-0.0275	-0.0969
	(0.0267)	(0.0320)	(0.216)
Profit	0.015^{**}	-0.033^{***}	-0.241^{***}
	(0.0067)	(0.008)	(0.052)
Highval_{t-1}*Profit	0.012	-0.008	-0.067
	(0.008)	(0.01)	(0.069)
Lowval_{t-1}*Profit	-0.001	-0.009	-0.085
	(0.007)	(0.009)	(0.059)
Constant	3.861^{***}	3.822^{***}	32.01^{***}
	(0.554)	(0.665)	(4.480)
Observations	330	330	330
R-squared	0.715	0.748	0.742

Standard errors in parentheses

$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 7: Regressions showing the effects of high and low realizations on investments in A and B and charity contributions

We see that low realizations of the unknown variable has a negative and significant effect on investment in A while the high realizations do not have any significant effect. This shows an asymmetry in behavior of the subjects from high and low values of the variable. This indicates an inter-temporal discouragement effect from the subjects. The extent of discouragement with low realizations of the variable is higher and more significant than the extent to which the subjects get encouraged from high realizations of the variable.

On the other hand we find that high realizations have a negative and significant impact on investment in B. With high realizations subjects allocate tokens away from project B, but that does not translate into a significant increase in project A. As a whole high values have a negative effect on charity earnings. Low values do not have any significant impact on charity earnings.

We see that the Risk+Bonus treatment has a negative impact on investment in A, investment in B as well as charity earnings. Moreover, there is an evidence of an intertemporal discouragement effect on investment in A where high values have a negative impact on charity while low values do not have any significant impact. Thus we can say that on the whole charity earning is affected negatively. Compared to the Bonus treatment introduction of risk in project A does not in any way make charity earning better.

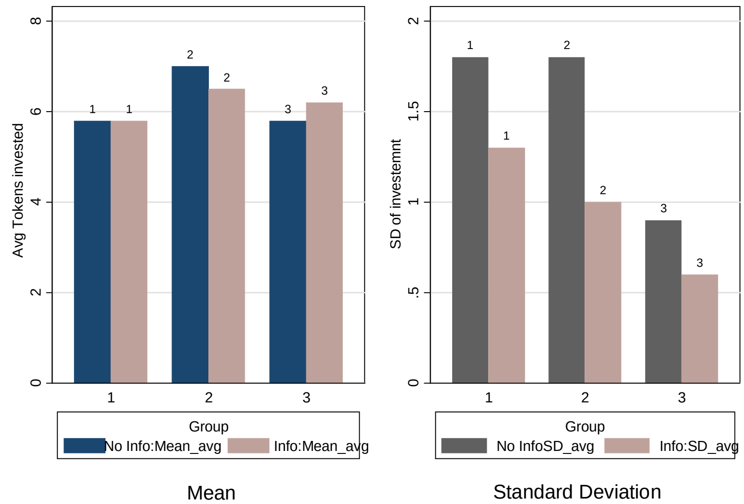

Figure 2: Average (left) and Standard Deviations (right) of tokens invested in proj. A in the last 4 periods of Bonus and Bonus+Info treatments for groups 1-3

6.4 Information

In this section we test the effect of information of past investment choices of group members on individual investment choices in A and B. We compare the effects of Information+Bonus treatment to that of the Bonus treatment. We start with the non-parametric Mann-Whitney test to find the effect of this treatment on investments in A and B, and the charity earnings. We find no observable difference between the treatments(all p-values equal to 1). We also want to test whether information on others' choices induce a clustering of investment choices around some group norm. Figures 1 and 2 show the mean and standard deviation of their choices for each group in this treatment as well as in the Bonus treatment. We include only the last 4 periods the figure. The figures show that although the average tokens invested remains the same across treatments the standard deviation of the choices decrease drastically in the Information+Bonus treatment compared to the Baseline treatment. This implies that with the information feedback on the group members' past choices, the subjects tend to cluster around the group average which they perceive as some measure of social norm. This might imply that on average information feedback has no significant effect on individual behaviors, but this can at least reduce the variance of the choices.[2]

For a detailed analysis of the effects of information on group behavior we run a set of panel regressions to see the effects of a group average on individual investment choices. We create a variable indicating the average of all the members in a group. Table 7 shows the effects of this group average on individual choices. We do not find any significant effect of group average on

[2] Appendix B shows kernel density estimates of the investment distributions in project B for Bonus and Bonus + Information treatments

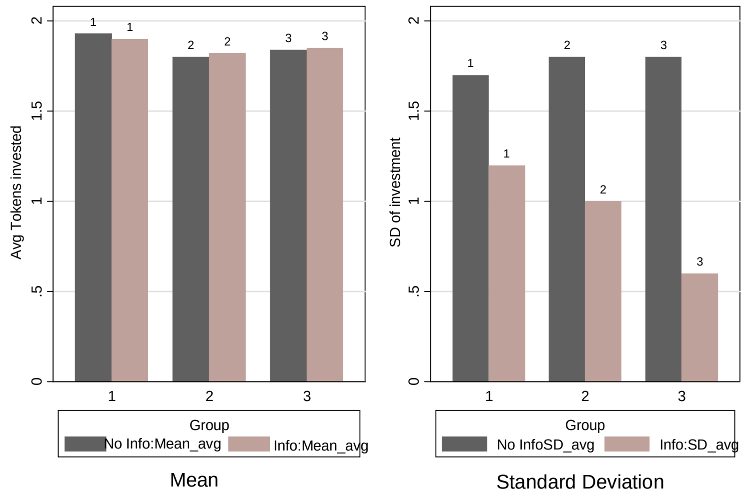

Figure 3: Average (left) and Standard Deviations (right) of tokens invested in proj. B in the last 4 periods of Bonus and Bonus+Info treatments for groups 1-3

individual choices. This can be due to two reasons– the first is the obvious reason the individual subjects behaved independently without any effect of the information they received about their group members. The other reason can be that there may be two opposing effects nullifying each other. Those who invested above the group average in the previous period reduce their investment in the current period, and those who invested below the average increase their investment in the current period. Table 8 shows regressions portraying the effects of group average on individual investments separately for individuals who are above and below the group average respectively. The variables $abovelagA$ and $belowlagA$ are two dummies indicating whether the individual's investment in A was above or below the group mean in the previous period. The negative and significant abovelagA and positive and significant belowlagA indicate that subjects who were above average decrease their investment while those who were below average increase their investment in A. The same is true for investment in B. These two significant opposing effects nullify each other. So the resultant effect of information on the overall investment choices is not significant. The result shows that information on group behavior changes the dynamics of individual investment choices, although it does not affect the aggregate investment.

	InvestA	**InvestB**
avgA_lag	−0.0815	
	(0.163)	
Period	0.00812	0.00275
	(0.0144)	(0.0235)
avgB_lag		−0.00754
		(0.163)
Constant	6.569***	2.091***
	(0.976)	(0.582)
Observations	225	225
R-squared	0.319	0.667

Standard errors in parentheses
$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 8: Regressions showing the effects of group average on aggregate investment

	InvestA	**InvestB**
abovelagA	−0.314*	
	(0.172)	
Period	0.00416	0.00776
	(0.0138)	(0.0230)
belowlagA	0.501***	
	(0.181)	
abovelagB		−1.104***
		(0.377)
belowlagB		0.814**
		(0.378)
Constant	6.279***	2.415***
	(0.273)	(0.440)
Observations	225	225
R-squared	0.428	0.680

Standard errors in parentheses
$^{***}p < 0.01, ^{**}p < 0.05, ^{*}p < 0.1$

Table 9: Regressions showing effect of group average on individual investments
(above and below the group average)

7 Conclusion

Performance-pay for teachers has raised a serious controversy in the education policy circle on whether we should incentivize teachers to perform their duties inside a classroom. Some of the studies in education literature claimed that if the incentives are based on the test scores of the students this might lead the teachers to skew their efforts towards teaching to the test rather than generating other higher-order learning skills among students, which they claim are necessary for the overall human capital formation of the society. Apart from this, some teachers' unions claim that the teachers are already motivated in their duties and an incentive might in fact demotivate them to perform their duties.

In order to address these potential issues we conducted an experiment where subjects faced similar incentives as that faced by the teachers. We find that the incentives have a substitution effect on the subjects' effort allocations leading to effort choices that generate more to their personal earnings substituting away from effort provision in the charitable task. Apart from this we do not find evidence of cross-crowding out of intrinsic motivation towards pro-social behaviors. We also find that non-deterministic earnings do not have any effect to decrease the power of incentives. Information feedback on the past choices of group members leads to a convergence of their choices around the group average, although in aggregate average effort choices remain the same.

Our results shed light on the broad perspectives of teachers' wage policy. We have shown that under similar financial incentives, there might exist certain adverse effects on individual effort allocations. This might create some potential problem for teachers' effort choices inside a classroom as external incentive might substitute away efforts towards teaching to the test thus sacrificing other aspects of teaching like generation of higher-order learning skills among students. These issues are in line with the standard substitution effects that the theory predicts and claims in some of the studies in the education literature. However the claims of the teachers' unions about cross crowding-out of intrinsic motivation cannot be justified from our data. Thus the education policymakers need to be aware of the potential issues that might exist while designing the wage policies for teachers. There should be a balance in the incentive structure so that these other aspects of teaching like generating higher-order skills among students are also not discouraged.

References

[1] Akerlof, G. (1980). A Theory of Social Customs, of Which Unemployment May be One Consequence. *Quarterly Journal of Economics*, 94(4), 749-775.

[2] Andreoni, J. (1993). An Experimental Test of the Public Goods Crowding-Out Hypothesis. *American Economic Review*, 83(5), 1317-1327.

[3] Andreoni, J. & D. Bernheim (2009). Social Image and the 50-50 Norm: A Theoretical and Experimantal Analysis of Audience Effects. *Econometrica*, 77(5), 1607-1636.

[4] Atkinson, A., S. Burgess, B. Croxon, P. Gregg, C. Propper & H. Slater, D. Wilson (2009). Evaluating the Impact of Performance Related Pay for Teachers in England. *Labor Economics*, 16(3), 251-261.

[5] Benabou, R. & J. Tirole (2003). Intrinsic & Extrinsic Motivation. *Review of Economic Studies*, 70, 489-520

[6] Bruggen, A. & F. Moers (2007). The Role of Financial Incentives and Social Incentives in a Multi-task Setting. *Journal of Management Accounting Research*, 19, 25-50.

[7] Cameron, J., W. Pierce, K. Banko & A. Gear (2005). Achievement-Based Rewards and Intrinsic Motivation: A Test of Cognitive Mediators. *Journal of Educational Psychology*, 97(4), 641-655

[8] Delisio, E.R. (2003a). Pay for Performance: What Are the Issues? *Education World.*

[9] Earl, L., Levin, B., Leithwood, K., Fullan, M., Watson, N., Torrance, N., Jantzi, D., Mascall, B., & Volante, L. (2003). *England's National Literacy and Numeracy Strategies: Final report of the external evaluation of the implementation of the strategies.* Department of Education and Employment, England.

[10] Fehr, E. & U. Fischbacher (2002). Why Social Preferences Matter. The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives. *Economic Journal*, C1-C33.

[11] Fehr, E. & K. Schmidt (2004). Fairness and Incentives in a Multi-task Proncipal-Agent Model. *Scandinavian Journal of Economics*, 106(3), 453-474.

[12] Fischbacher, U. (2007). z-Tree: Zurich Toolbox For Readymade Economic Experiments. *Experimental Economics*, 10(2), 171-178.

[13] Frey, B. (1997). From the Price to the Crowding Out Effect, *Swiss Journal of Economics and Statistics*, 133, 325-350.

[14] Glewwe, P, N. Ilias & M. Kremer (2010). Teacher Incetives. *American Economic Journal(Microeconomics)*, 2(3), 205-227.

[15] Gneezy, U. & A. Rustichini (2000). Pay Enough or Don't Pay At All. *Quarterly Journal of Economics*, 115(3), 791-810.

[16] Goodman, S. & L. Turner (2013). Teacher Incentive Pay and Educational Outcomes: Evidence from the New York City Bonus Program. *Journal of Labor Economics*, 31(2), 409-420.

[17] Hannaway, J. (1991). Higher-order Thinking, Job Design and Incentives: An Analysis and Proposal. *American Educational Research Journal*, 29(1), 3-21.

[18] Hanushek, E. (2003). A More Complete Picture of School Resource Policies. *Review of Educational Research*, 66(3), 397-409.

[19] Hanushek, E. (2003). The Failure of Input-based Schooling Policies. *The Economic Journal*, 113, 64-98.

[20] Hanushek, E.& J. Kain & S. Rivkin. (2005). Teachers, Schools and Academic Achievement. *Econometrica* 73(20), 417-458.

[21] Hanushek, E. (2011). The Economic Value of Higher Teacher Quality, *Economics of Education Review*, 30, 466-479.

[22] Harbison, R.W., Hanushek, E.A. (1992). Educational Performance of the Poor: Lessons from Rural Northeast Brazil. *Oxford University Press*, New York.

[23] Harvey, J. (2005). Why did you do that? An Economic Examination of the Effects of Extrinsic Compensation on Intrinsic Motivation and Performance. *Journal of Economic Psychology*, 26, 546-566.

[24] Herman, J. L. (1992). What research tells us about good assessment. *Educational Leadership*, 49(8), 74-78.

[25] Holmstrom, B. & P. Milgrom (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, & Organization,* 29(4), 24-52.

[26] Hoppe, E. & D. Kusterer (2011). Conflicting Task and Moral Hazard: Theory and Experimental Evidence. *Europian Economic Review*, 55(8), 1094-1108.

[27] Hudson, S. (2010). The Effects of Performance-basedTeacher Pay on Student Achievement. *SIEPR Discussion Paper* No. 09-023.

[28] Kandori, M. (1992). Social Norm and Community Enforcement. *Review of Economic Studies*, 59(1), 63-80.

[29] Kane, T. & D. Staiger. (2008). Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation, *NBER Working Papers* 1407, National Bureau of Economic Research, Inc.

[30] Kunz A. & D. Pfaff. (2002). Agency theory, performance evaluation and the hypothetical construct of intrinsic motivation. *Accounting, Organizations and Society*, 27, 275-295

[31] Lavy V. (2007). Using performance-based pay to improve the quality of teachers. *NCBI Working Papers*, 17(1):87-109.

[32] Lindbeck, A., S. Nyberg & J. Weibull. (1999). Social Norms and Economic Incentives in the Welfare State. *Quarterly Journal of Economics*, 114(1), 1-35.

[33] Lockheed, M. & A. Verspoor (1991). Improving Primary Education in Developing Countries, *Oxford University Press for the World Bank*.

[34] Mago, S., A. Samak & R. Seremeta. Facing Your Opponents: Social Identification and Information Feedback in Contests. *Journal of Conflict Resolution*, forthcoming.

[35] Muralidharan, K. & V. Sundaraman (2010). The Impact of Diagnostic Feedback to Teachers on Student Learning: Experimental Evidence from India. *Economic Journal*, 120(546), F187-F203

[36] Payne, D. A. (2003). *Applied Educational Assessment (2nd ed.)*. Toronto: Wadsworth Group.

[37] Popham, W. J. (2001). Teaching to the test. *Educational Leadership*, 58(6), 16-20.

[38] Rivkin, S., Hanushek, E. & Kain, J. (2005). Teachers, Schools and Academic Achievement. *Econometrica,* 73(2), 417-458.

[39] Rockoff, J. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data, *American Economic Review*, 94(2), 247-252.

[40] Ryan, R. & E. Deci (2000). When Rewards Compete with Nature: the Undermining of Intrinsic Motivation and Self-regulation. C. Sansone & J. Harachiewicz (Eds.) *Intrinsic and Extrinsic Motivation: The Search for Optimal Motivation and Performance.* (14-54), San Diego Academic Press.

[41] Solow, R. (1992). The Labor Market as a Social Institution. *Journal of Economic Issues.* 26(4), 1301-1303.

[42] Vansteenkiste, M., W. Lens & E. Deci (2006). Intrinsic vs Extrinsic Goal Contents in Self-determination Theory. *Educational Psychology*, 41(1), 19-31

[43] Woessmann, L. (2014). Cross-country Evidence on Teacher Performance Pay. *Program on Education Policy and Governance Working Papers Series.*

[44] Zanella, B. (1998). Verdrängung von Reziprozität durch Anreizverträge. *Mimeo*, Institute of Empirical Economic Research, University of Zurich.

......

Appendix A

Let the effort levels for the two tasks be e_1 and e_2 respectively where $e_1,\ e_2 \in [\underline{e}, \overline{e}]$. and $\leq \underline{e}, \overline{e}, \leq 1$. Let the cost function of the agent for the two effort levels be

$$c(e_1, e_2) = \frac{1}{2}c_1 e_1^2 + \frac{1}{2}c_2 e_2^2 + \gamma e_1 e_2$$

where $c_1, c_2, \gamma \geq 0$. We assume that $c_1 \leq c_2$. Let $\theta_1 e_1$ and $\theta_2 e_2$ be the inherent utilities that the agent derives from performing the tasks, with $\theta_1, \theta_2 \geq 0$.

When both the e_1 and e_2 are perfectly observable by the principal, the overall utility for the agents is,

$$U(e_1, e_2) = \theta_1 e_1 + \theta_2 e_2 - \frac{1}{2}c_1 e_1^2 - \frac{1}{2}c_2 e_2^2 - \gamma e_1 e_2.$$

Since the efforts are perfectly observable, the solution is the first best. Let $f_1 e_1$ and $f_2 e_2$ be the contributions of the different tasks to the overall welfare of the society with $f_1 < f_2$.

The welfare function of the society is

$$W(e_1, e_2) = f_1 e_1 + f_2 e_2 + \theta_1 e_1 + \theta_2 e_2 - \frac{1}{2}c_1 e_1^2 - \frac{1}{2}c_2 e_2^2 - \gamma e_1 e_2.$$

The first order conditions to maximize the welfare function give the first best solutions of the effort levels.

$$e_1^{FB} = \frac{\gamma(f_2 + \theta_2) + c_2(f_1 + \theta_1)}{c_1 c_2 - \gamma^2}.$$

and

$$e_2^{FB} = \frac{\gamma(f_1 + \theta_1) + c_1(f_2 + \theta_2)}{c_1 c_2 - \gamma^2}.$$

The second order condition for a local maximum gives $c_1 c_2 > \gamma^2$. Thus $\gamma < \sqrt{c_1 c_2}$ is the feasible region where γ lies in order to satisfy the second order condition. Note that when $\gamma = \sqrt{c_1 c_2}$, the two effort levels are perfectly substitutable, and interior solution does not exist.

These are the values of e_1 and e_2 in the case of interior solutions. If we allow for corner solutions, for any $e_j = \overline{e}$, $e_i = \frac{f_i + \theta_i - \gamma\overline{e}}{c_i}$, $i, j = \{1, 2\}, i \neq j$. The same is true for $e_j = \underline{e}$. Thus the set of solutions become :

$$\widehat{e_{1,FB}} = \begin{cases} \frac{f_1 + \theta_1 - \gamma\underline{e}}{c_1} \text{ if } e_2^{FB} \leq \underline{e} \\ e_1^{FB} \text{ if } \underline{e} \leq e_2^{FB} \leq \overline{e} \\ \frac{f_1 + \theta_1 - \gamma\overline{e}}{c_1} \text{ if } e_2^{FB} \geq \overline{e} \end{cases}$$

and

$$\widehat{e_{2,FB}} = \begin{cases} \frac{f_2+\theta_2-\gamma \underline{e}}{c_2} \text{ if } e_1^{FB} \leq \underline{e} \\ \widehat{e_2} \text{ if } \underline{e} \leq e_1^{FB} \leq \overline{e} \\ \frac{f_2+\theta_2-\gamma \overline{e}}{c_2} \text{ if } e_1^{FB} \geq \overline{e} \end{cases}$$

Allowing for corner solutions, since there are upper and lower limits to both the effort levels, the complete set of first best solutions becomes :

$$e^*_{1,FB} = \max\{\min\{\widehat{e_{1,FB}}, \overline{e}\}, \underline{e}\}.$$

and

$$e^*_{2,FB} = \max\{\min\{\widehat{e_{2,FB}}, \overline{e}\}, \underline{e}\}.$$

In the case where the efforts are not observable, let the principal offer a bonus rate b to incentivize effort from the agents on the basis of the score function denoted by

$$P = g_1 e_1 + g_2 e_2.$$

The agents' problem is to maximize :

$$Max_{e_1,e_2} s + b(g_1 e_1 + g_2 e_2) + \theta_1 e_1 + \theta_2 e_2 - \frac{1}{2} c_1 e_1^2 - \frac{1}{2} c_2 e_2^2 - \gamma e_1 e_2$$

where s is a fixed wage independent of the effort choices. The first order condition gives the following effort levels:

$$\widehat{e_1} = \frac{(c_2 g_1 - \gamma g_2)b + c_2\theta_1 - \gamma\theta_2}{c_1 c_2 - \gamma^2}$$

and

$$\widehat{e_2} = \frac{(c_1 g_2 - \gamma g_1)b + c_1\theta_2 - \gamma\theta_1}{c_1 c_2 - \gamma^2}.$$

These are the values of e_1 and e_2 in the case of interior solutions. Again, as in the case of first best solutions, the same logic applies here if we allow for corner solutions. For any $e_j = \overline{e}$, $e_i = \frac{bg_i+\theta_i-\gamma\overline{e}}{c_i}$, $i, j = \{1, 2\}, i \neq j$. The same is true for $e_j = \underline{e}$. Thus the set of solutions become :

$$e_1^{SB} = \begin{cases} \frac{bg_1+\theta_1-\gamma \underline{e}}{c_1} \text{ if } \widehat{e_2} \leq \underline{e} \\ \widehat{e_1} \text{ if } \underline{e} \leq \widehat{e_2} \leq \overline{e} \\ \frac{bg_1+\theta_1-\gamma \overline{e}}{c_1} \text{ if } \widehat{e_2} \geq \overline{e} \end{cases}$$

and

$$e_2^{SB} = \begin{cases} \frac{bg_2+\theta_2-\gamma \underline{e}}{c_2} \text{ if } \widehat{e_1} \leq \underline{e} \\ \widehat{e_2} \text{ if } \underline{e} \leq \widehat{e_1} \leq \overline{e} \\ \frac{bg_2+\theta_2-\gamma \overline{e}}{c_2} \text{ if } \widehat{e_1} \geq \overline{e} \end{cases}$$

Again, the complete set of second best condition becomes :

$$
\begin{aligned}
e^*_{1,SB} &= \max\{\min\{e_1^{SB}, \overline{e}\}, \underline{e}\} \\
e^*_{2,SB} &= \max\{\min\{e_2^{SB}, \overline{e}\}, \underline{e}\}.
\end{aligned}
$$

Thus $e^*_{1,SB}$ and $e^*_{2,SB}$ constitute the incentive compatibility constraint (ICC) for the agent. The participation constraint (PC) for the agent is

$$
s + b(g_1e_1 + g_2e_2) + \theta_1e_1 + \theta_2e_2 - \frac{1}{2}c_1e_1^2 - \frac{1}{2}c_2e_2^2 - \gamma e_1e_2 \geq \underline{u}
$$

where $\underline{u}$ is some reservation utility level. Along with this, there is also a limited liability constraint (LLC) : $s + b(g_1e_1 + g_2e_2) \geq s$ and $s \geq s$. We can assume that the value of $\underline{u}$ is so low that the participation constraint does not bind. So we can ignore the participation constraint, and the relevant binding constraints are the ICC and the LLC.

The principal's optimization problem is thus :

$$
Max_b f_1e_1 + f_2e_2 - s - b(g_1e_1 + g_2e_2)
$$

$s.t.$ 1) *Incentive compatibility constraint* : $e_1 = \max\{\min\{e_1^{SB}, \overline{e}\}, \underline{e}\}$ and $e_2 = \max\{\min\{e_2^{SB}, \overline{e}\}, \underline{e}\}$.

2) *Limited liability constraint* : $s + b(g_1e_1 + g_2e_2) \geq s$ and $s \geq s$.

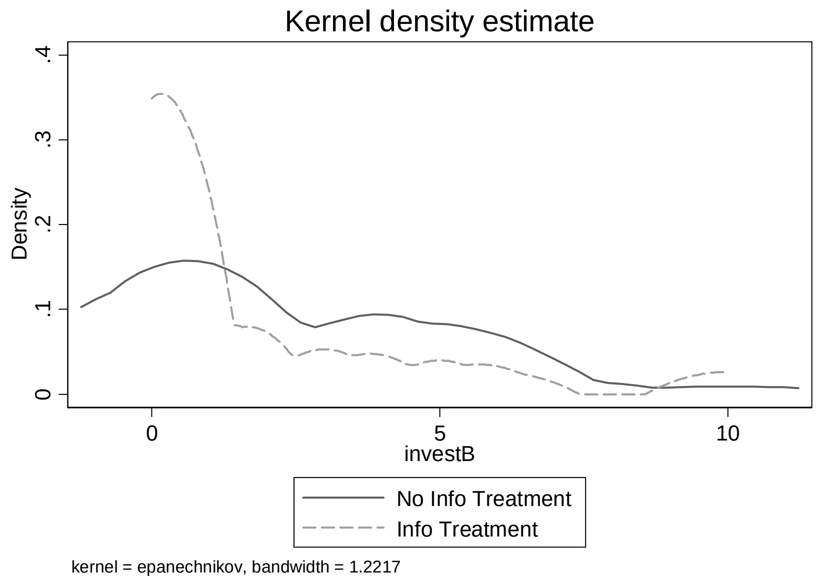

Figure 4: Kernel Density Estimate of the distributions of investments in proj. B for Bonus and Bonus+Info treatments