Incentives for teachers: A theoretical framework and a structural model

Ronak Jain*
Harvard University

Aug 2018

Abstract

This paper provides structural analysis of teacher behaviour under the provision of individual and group monetary incentive schemes using data from a randomised control trial conducted in India. I develop a theoretical principal-agent model explicitly allowing for both intrinsic motivation and a crowding out effect of explicit performance pay on intrinsic incentives. Theoretically, I show that intrinsic motivation and extrinsic motivation through performance pay are substitutes under the optimal incentive scheme in the presence of motivation crowding out. Applying the theoretical framework, I identify the sets of teacher heterogeneity consistent with each effort choice in the treatment and control group schools. Using the barycenters of these sets, I find that increasing the effort of the average teacher who did not change his/her effort under the experimental incentive payment of Rs. 500, would require increasing the bonus payment to Rs. 730 under the individual incentive scheme. The corresponding figure lies between Rs. 720 - Rs. 1020 depending on the reported group cooperation level under the group incentive scheme.

Keywords: performance pay, group incentives, intrinsic motivation, motivation crowding out

*This paper is based on my MPhil thesis at the University of Oxford. I am very grateful to my supervisors, Ian Crawford and Margaret Meyer, for their guidance and support for this project. I thank Pramila Krishnan and Abhijeet Singh for encouragement, and Rossa O'Keeffe-O'Donovan, Ian Jewitt, Simon Quinn, and seminar participants at The Busara Center for Behavioral Economics for helpful suggestions and discussions. All errors are my own. Comments are very welcome at rjain@g.harvard.edu

1 Introduction

Given the importance of the role of a teacher in a child's education, it is important to have an understanding of how incentives might affect a teacher's motivation level and student learning outcomes. Indeed, the significance of this is even greater for developing economies where teacher absenteeism and low teacher effort is prevalent (Chaudhury et al., 2006, Kremer et al., 2005, EFA Global Monitoring Report, 2014). Whilst much of the recent experimental literature has focussed on measuring the reduced form effect of using incentives to help redress the situation, we still lack an understanding of an intermediate step - how explicit incentives interact with certain teacher characteristics which are key determinants of a teacher's responsiveness to incentives. An understanding of these interaction effects is valuable as it can enable us to better target the use of performance pay in this context.

A number of studies have highlighted that the potential unintended consequences of incentive provision, such as crowding out of intrinsic motivation or signalling distrust on the part of the principal, can make incentive provision counter productive in different contexts (Frey and Oberholzer-Gee, 1997, Gneezy and Rustichini, 2000, Falk and Kosfeld, 2006, Ariely et al., 2009, Fuster and Meier, 2010, Gneezy et al., 2011). The magnitude of the incentive and how the agent perceives a particular form of incentive also plays an important role in avoiding perverse outcomes (Gneezy and Rustichini, 2000). Whilst the existing theoretical literature sheds light on factors such as multi-tasking (Holmstrom and Milgrom, 1991) and the necessity to base incentive pay on an imperfect but easily quantifiable performance measure indicative of the effort (Baker, 1992, Baker, 2000, Neal, 2011), which make it difficult to unequivocally predict the effectiveness of use of performance pay for teachers, there is a dearth of theoretical literature that formally models the possible unintended effects noted in the preceding.

In particular, it is plausible that since teaching is perceived as an activity that makes a valuable contribution to the society, using performance related pay for teachers might risk displacing their intrinsic motivation that stems from their self-image concerns or feelings of warm glow. Addressing this lacuna in the theoretical literature, to the best of my knowledge, this paper is among the first to develop a theoretical principal-agent model incorporating both intrinsic motivation and a crowding-out effect of explicit performance pay on intrinsic incentives.[1]

Experimental evaluations on the use of financial incentives for teachers have found varying degrees of success (Neal, 2011). For instance, Duflo et al.(2012) found that monitoring combined with financial incentives reduced teacher absenteeism by 21 percentage points in rural India and increased student test scores by 0.17 standard deviations. Muralidharan and Sundararaman (2011) also find that performance based pay for government teachers in India led to an increase in student math and language scores by 0.27 and 0.17 standard deviations respectively. However, Barrera-Osorio and Raju (2017) find that giving bonuses to teachers in Pakistan did not lead to improvement in test scores though enrolment and exam participation of students increased. Whilst the results of

[1] Whilst Besley and Ghatak (2005 & 2008) model public servants as likely to differ systematically from those who work in the private sector, for example, in terms of their motivation level or image concerns, their model does not consider how they might also additionally respond differently to extrinsic incentives.

Glewwe et al.(2010) involving Kenyan primary school teachers suggest a positive impact of performance linked pay on incentivised subjects, there was no improvement in teacher attendance. The heterogeneity of findings of these experimental studies serve to highlight that a consideration of how and why the responsiveness of teachers to incentives may differ across contexts merits an investigation.

Importantly, although experimental studies can help to judge the end result of providing incentives, an understanding of the precise mechanism through which the effect is brought about is left wanting by relying solely on this method to approach the question. Philosopher Immanuel Kant reasoned that "concepts without percepts are empty, percepts without concepts are blind", arguing that concepts, such as space and time, are a prerequisite for us to have intelligible perceptual experience.[2] Analogously, I believe that a theoretical model enables us to interpret the mechanisms driving the reduced form results obtained from experimental studies and experimental analysis acts as a check for predictions of the theoretical model. Exploiting this synergy, I use a structural modelling approach to analyse teacher response to incentive pay, applying the theoretical framework to analyse the findings of a randomised control trial in India that provided financial incentives to government primary school teachers based on improvement in student achievement (Muralidharan and Sundararaman, 2011). Using this approach offers the opportunity to conduct the counterfactual analysis of the amount by which the incentive pay would need to increase in order to raise teacher effort even further (Todd and Wolpin, 2010, Bandiera et al., 2011, Keane et al., 2011, Low and Meghir, 2017). It is in this respect that the paper seeks to contribute to the existing literature on the use of incentive pay for teachers.

This experimental study was conducted in schools in Andhra Pradesh, India.[3] The experiment had one treatment group of schools where teachers received bonuses based on the performance of their students (teacher-level incentives) and another group of schools where teachers received bonus pay based on the performance of all the students in their school (group-level incentives). Explicitly modelling important dimensions of teacher heterogeneity, the paper first theoretically investigates: (1) how does teacher behaviour change when given individual-level incentives; (2) how do teachers respond when given group incentives; and (3) the optimal incentive pay under both these incentive schemes.

Theoretically, I find that under the individual incentive scheme, the optimal incentive pay is decreasing in a teacher's intrinsic motivation and the extent to which a teacher experiences motivation crowding out. When providing group-level incentives, I find that the optimal incentive pay is decreasing in a teacher's intrinsic motivation and if the teachers are similar in terms of their ability and the extent to which they experience motivation crowding out, then the optimal incentive pay is also decreasing in their motivation crowding out. Hence, when providing individual-level or group incentives, extrinsic and intrinsic motivation act as substitutes for the principal under the optimal incentive contract. Interestingly, for a given incentive payment schedule, if a teacher experiences high motivation crowding out but is also highly altruistic towards other teacher's earnings from the incentive scheme, then providing group-level incentives could dominate individual-level

[2] Kant, Immanuel. Critique of pure reason. *Cambridge University Press*, 1998.

[3] The next section describes the experiment in more detail. Please refer to the Muralidharan and Sundararaman (2011) paper for an overview of the characteristics of this state which made it an ideal setting for the experiment.

incentives for incentivising effort.

In structural analysis for the experiment, I refrain from making any untestable parametric assumptions about the nature of the underlying teacher heterogeneity or treating it simply as a multivariate error term. Instead, I use a set identification strategy for empirical analysis. I find that the density of intrinsic motivation in the control schools is bimodal with teachers either having very low or medium levels of intrinsic motivation. I also identify the set for teacher heterogeneity consistent with each effort choice in the treatment schools. Using these sets, I then derive the bounds for the incentive payment needed to raise teachers' efforts. I find that the upper bound for the increase in incentive payment needed depends on the underlying teacher heterogeneity; in particular, it is decreasing in the teachers' ability and increasing in teachers' motivation crowding out. I also show that it is comparatively less costly to incentivise all teachers to increase their effort by one if their initial effort choice was unchanged under the experiment than if the teachers did increase their effort under the experimental scheme. Furthermore, I find that the cost of increasing all teachers' efforts from that chosen under the experiment is increasing in the magnitude of the increase in effort required.

In particular, I find that increasing the effort of the average teacher who did not change his/her effort under the incentive scheme, would require increasing the incentive bonus to Rs. 730 for the individual-level incentives whilst the figure lies between Rs. 720 - Rs. 1020 depending on the reported group cooperation level among teachers when providing group incentives. For increasing the change in effort of the teachers not choosing the maximal level of effort under the original payment scheme, we would need an incentive payment of at least Rs. 1020 under individual-level incentives and at-least Rs. 850 - Rs. 1020 under the group incentive payment scheme depending on the reported group cooperation level among teachers.

The rest of the paper proceeds as follows. Section 2 briefly outlines the design and findings of the experimental incentive program. Section 3 lays out the theoretical framework for teacher-level and group-level incentives. Section 4 presents the structural model and estimation results. Section 5 provides a discussion of the findings and section 6 concludes.

2 A brief overview of the experiment

The experiment gave primary school teachers incentives based on either their own students' test performance (individual-level incentives) or the test performance of all the students in their school (group incentives). The randomised control trial involved 300 government-run primary schools in the Indian state of Andhra Pradesh over the time period 2005-07.[4] 100 schools were allocated to the treatment with individual-level incentives, 100 schools to group-level incentives and 100 schools were in the control group. Randomization was carried out at the level of a school in a mandal, which typically

[4] The communication to teachers with respect to the length of the program was that the program would continue as long as the government continued to support the project. The expectation conveyed to teachers during the first year was that the program was likely to continue but was not guaranteed to do so.

has 25 villages.[5] Teachers in the government-run schools are employed by the state and their salary is determined according to their experience and rank. There is no existing performance-related pay component in their compensation structure.

Students were assessed in math and language using two tests which assessed the materials covered in the previous year and the current school year. Teachers were incentivised with the following piece-wise linear bonus schedule for the first year:

$$Bonus = \begin{cases} Rs.\ 500 * (percent\ gain\ in\ average\ scores - 5) & if\ gain > 5\% \\ 0 & otherwise \end{cases}$$

A threshold of 5% was used in the first year to allow for the possibility of students not having some of the taught material fresh in their minds given that the baseline tests were conducted following the summer holidays. In the second year of the experiment, the bonus schedule was linear with a threshold of zero as the end of year tests in the first year formed the baseline tests for the second year. In the case of individual-level incentives the performance measure used was percent gain in average test score of a teacher's own students whilst in the case of group incentives, the performance measure used was percent improvement in average school-level scores. This bonus payment slope of Rs. 500 amounts to $1/16^{th}$ of a typical teacher's salary.

The experiment found that student test scores in math and language were higher by 0.27 and 0.17 standard deviations respectively in incentive schools. The authors also find that teacher-level incentives outperformed group-level incentives at the end of two years whilst they performed equally well in the first year of the experiment. The experiment involved several rounds of unannounced visits for recording teacher attendance and activity. Following the end-of-year tests and before the results were released, teachers were interviewed about their response to the incentive scheme. In particular they were asked whether they gave any special preparation for their students, assigned extra homework or classwork and focused on weaker students. I first lay out the theoretical framework which can be used to understand the set-up of the experiment in the next section and subsequently use the self-reported data on teacher behaviour to recover sets of the structural parameters.

3 Theoretical Framework

3.1 Provision of individual-level incentives with intrinsically motivated agents

Individuals vary considerably in their level of intrinsic motivation. A number of studies have noted the importance of intrinsic motivation in influencing the behaviour of workers

[5] The details of the randomization procedure can be found in Muralidharan and Sundararaman (2011). Briefly, 5 districts were first chosen in Andhra Pradhesh. In each of these districts, one division was randomly selected and 10 mandals were then randomly selected per division. 10 schools per mandal were then randomly selected using probability proportional to enrolment. 2 schools in each mandal were assigned to individual-level incentives treatment, group incentives treatment and control group respectively.

and have highlighted the possible danger of monetary incentives displacing this.[6] In light of these studies, I model teachers as motivated agents and going beyond the existing theoretical literature on incentives, I explicitly allow for the possibility of a motivation crowding out effect in the utility function.

In order to focus on illustrating how the effectiveness of providing incentives varies with the level of one's intrinsic motivation, I use a stylised static model. Rather than restricting the analysis strictly to the experimental setting, the theoretical framework outlined here examines the use of performance related pay based on student improvement more broadly. The government (principal) and teachers (agents) are risk neutral. Effort e, chosen by the teacher, is unobservable directly by the government and thereby, non-contractible.[7] Instead, the government can observe student test scores in a subject $S(\mathbf{e}, \mu, \iota, \eta)$ which are influenced by cumulative effect of efforts of the student's current and former teachers in that subject ($\mathbf{e}$), a student's unobservable and time-invariant ability μ, a school fixed effect capturing school inputs ω and an additively separable mean zero idiosyncratic shock η. Hence the test score of a student j of a teacher i at a school s in a year t can be expressed as

$$S_{ijts}(\mathbf{e}, \mu, \iota, \eta) = \sum_{i_t} g_{i_t} e_{i_t} + \mu_j + \omega_s + \eta_{jt}$$

where g_{i_t} is the marginal product of a teacher i's effort on the student's test score. This can also be interpreted as the teacher's 'ability' to convert effort into student test scores.

The teacher is offered a linear contract $w = \alpha + \beta \tilde{S_{it}}$ based on the improvement in the average test score of his/her students denoted by[8]

$$\tilde{S}_{it}(e_{i_t}, \eta) = \frac{\sum_j (S_{ijt} - S_{ijt-1})}{m_i}$$

where m_i denotes the number of students of a teacher. Note that incentivising teachers based on the change in a student's test score over the academic year prevents the student's ability from driving a teacher's compensation. Hence μ_j no longer appears as an argument in the $\tilde{S_{it}}(e_{i_t}, \eta)$ function. Similarly, differencing eliminates all student, teacher and school specific time invariant heterogeneity, provided these enter additively in the production function for student test scores, and hence we can ignore these in our analysis herein.

The government cares about developing the human capital of students $P(e, \varepsilon)$, which in turn is a function of a teacher's effort and an additively separable mean-zero idiosyncratic shock ε. The government's payoff can therefore be represented as

[6] See for example Frey and Oberholzer-Gee (1997), Gneezy and Rustichini (2000), Ariely et al.(2009) and Gneezy et al. (2011)

[7] This assumption is a realistic one to make in contexts where government plays the role of the principal and is unable to monitor the employees of the state.

[8] The subscript s for school is suppressed for ease of notation.

$$u_G = P(e, \varepsilon) - (\alpha + \beta \tilde{S}_{it}(e, \eta))$$

which is the benefit that the government derives from the human capital formation of students less the cost it has to bear to pay and incentivise a teacher. The utility of a teacher is

$$u_T = \alpha + \beta \tilde{S}_{it}(e, \eta) + \phi(e, \beta) - c(e)$$

where $\phi(e, \beta)$ is the motivation function.[9][10]

The motivation function captures two different and contrasting effects: (1) a higher marginal benefit of exerting effort representing the intrinsic motivation of a teacher and (2) a motivation crowding out effect that a teacher experiences due to the provision of explicit monetary incentives. To see this more explicitly, let $\phi(e, \beta) = \kappa(e) + \varsigma(e, \beta)$. Here the term $\kappa(e)$ captures the first effect with $\kappa'(e) > 0$, that is, all else equal a motivated agent receives a higher marginal benefit of exerting effort. This may reflect the psychological satisfaction obtained from contributing to a public good or doing one's duty. The second term $\varsigma(e, \beta)$ captures the second effect with $\varsigma_{12}(e, \beta) < 0$, that is, the agent's marginal utility from exerting effort is decreasing in the monetary incentives provided.[11] One may also find it useful to think of intrinsic motivation, and hence the dependence of the motivation function on effort, being derived from the fact that a teacher cares about their student's capital formation to some extent. In this case, the motivation function will depend on e indirectly: $\phi(P(e, \varepsilon), \beta))$. For expositional clarity and ease, I stick to using e as a direct argument of the motivation function.

It is reasonable to assume that $\phi_{11}(e, \beta) \leqslant 0$ so that there no increasing returns of exerting effort on one's utility. Costs of exerting effort are assumed to be increasing and convex so that $c'(e) > 0$ and $c''(e) > 0$. Government's payoff is increasing in the teacher's effort but the marginal benefit of effort on the payoff is non-increasing in e so that $P_1(e, \varepsilon) > 0$ and $P_{11}(e, \varepsilon) \leqslant 0$. Similarly, student test scores are increasing in a teacher's effort so that $S_1(e, \eta) > 0$[12] but there are non-increasing returns to effort with $S_{11}(e, \eta) \leqslant 0$. The government's payoff only depends on the teacher's effort and not the specific teacher it is matched with.

I have abstracted from a multi-tasking model and I take effort to be uni-dimensional. This is a plausible assumption to make in a context, such as the present one, where that teachers are rewarded based on the increase in test scores of their students. This is because the different tasks of a teacher such as marking homework, classroom teaching and exam

[9] For ease of exposition the function $\phi(\beta, e)$ is assumed to be continuous and twice differentiable with monotonic and continuous partial derivatives.

[10] Since we are only considering the principal as interacting with one agent here and because the problem is static, for notational ease, hereafter I leave the index $\{it\}$ on the performance measure $\tilde{S}_{it}$, and $\{i\}$ on the teacher's effort e and marginal product of teacher's effort g, implicit in this subsection.

[11] This also equates to assuming that the extrinsic motivation through β and intrinsic motivation represented by $\phi_1(e, \beta)$ are substitutes in the agent's motivation function.

[12] This is equivalent to $g \geqslant 0$.

preparation will all contribute to a student's test score performance, that is, are likely to be complements in the production technology of test scores. A teacher is therefore likely to increase effort in one or more of these tasks rather than substitute between them in response to performance related pay based on an increase in student test scores. Hence an increase in one or more of these can be captured as an increase in effort. A further justification for this follows from the consideration that teachers may be unaware of the exact technology that determines student test scores and consequentially, are likely to increase effort in one or more of these rather than substitute one task for another.[13]

The timing of events is as follows.

1. The government stipulates the contract including the performance related pay component.
2. A teacher accepts or rejects the contract. If a teacher rejects the contract, he/she gets the reservation utility $\overline{u}$ from an outside option. If the contract is accepted, a teacher maximises his/her expected utility choosing their privately-observed effort subject to the participation constraint $E[u_T|e] \geqslant \overline{u}$.
3. Random shocks (ε, η) are realised but unobserved by the government. Both the government and the teacher can observe the student test scores and teachers are paid in accordance with the stipulated contract.

To see the implications of including a motivation function, it is helpful to make a comparison with the standard principal-agent model. For this purpose, I refer to an agent with the motivation function included in the utility as the motivated agent and to an agent without the motivation function as the standard self-interested agent with a utility function $\alpha + \beta\tilde{S}(e, \eta) - c(e)$. The following results help to illustrate the difference.

Lemma 1. *Incentive compatibility dictates*[14]*:* $\beta\tilde{S}_1(e, \eta) + \phi_1(e, \beta) = c'(e)$

One can observe that if the intrinsic motivation of the agent is sufficiently large to outweigh the effect of motivation crowding out so that $\phi_1(e, \beta) > 0$, the effort exerted by this type of agent, for a given β, is higher than the standard self-interested agent whose incentive compatibility would dictate $\beta\tilde{S}_1(e, \eta) = c'(e)$.[15]

The agents also differ in their responsiveness to an increase in the intensity of incentives as the following corollary notes.

Corollary 1. *The motivated agent is less responsive to an increase in the intensity of incentives than a self-interested agent.*[16]

[13]Nonetheless, this assumption is innocuous to the implications of the model and can be relaxed, for instance, by making effort two-dimensional.

[14]The concavity of the objective function ensures an interior maximum. Hence the first order condition suffices to get the incentive compatible effort level. As the error in the technology for test scores is assumed to be additively separable and the agent is risk neutral, the expectations operator can be suppressed in the analysis here.

[15]This follows from the convexity of the cost function.

[16]Please refer to the Appendix for the formal proof of this and all subsequent results.

I now make some simplifying assumptions in order to shed light on key insights that can be derived from this framework. The cost function is $c(e) = \frac{1}{2}(e - e_0)^2$ for $e \geqslant e_0$ where e_0 is the teacher's effort last period. The teacher therefore experiences disutility from increasing effort above this level. If the teacher decreases his/her effort, then I assume that the disutility of effort is zero.[17] Let $P = fe + \varepsilon$ and the test score technology is as specified before. Hence $E[\tilde{S}(e, \eta)|e] = g\tilde{e}$ where $\tilde{e} = e - e_0$. $P_1(e, \varepsilon)$, the marginal benefit of effort on the principal's payoff is assumed to be independent of the level of e and therefore, $P_{11}(e, \varepsilon) = 0$.

To compare the optimal contract for a motivated agent to that of a self-interested agent in a simple and tractable form, I assume a linear motivation function $\phi(e, \beta) = ke - \gamma\beta g\tilde{e}$, where $\gamma \geq 0$ is a scale parameter denoting the extent to which there is a motivation crowding out effect of providing performance related pay.[18] This formulation of the motivation function captures the intrinsic motivation of a teacher via $\kappa'(e) = k$ where $k > 0$ and the motivation crowding out effect via $\phi_{12}(e, \beta) = -\gamma g$ so that $\phi_{12} < 0$. There are non-increasing returns to effort with $\phi_{11}(e, \beta) = 0$.

I focus on non-negative values of γ allowing for the possibility of a motivation crowding out effect. A value of $\gamma = 1$ would mean that there is a one-to-one crowding out effect of monetary incentives whilst a value of $\gamma > 1$ would mean that the motivation crowding out effect is so strong that the overall effect of providing monetary incentives on a teacher's utility is negative.[19] Since in these scenarios, the optimal incentive intensity is trivially zero, I focus my analysis for $\gamma \in [0, 1)$. Note that with this formulation $k = \gamma = 0$ would give a standard self-interested agent.

Teachers differ in terms of their marginal product of effort on raising test scores or 'ability' g, in the extent to which they experience motivation crowding out γ, and their intrinsic motivation k. These three dimensions constitute the heterogeneity among teachers. I assume here that teachers know their ability g and preferences k and γ. The diagram below helps to illustrate how the marginal utility of effort for a given β differs according to the assumptions made on the parameters governing intrinsic motivation and the motivation crowding out effect.

The case with $k = \gamma = 0$ in red gives the marginal benefit for a standard agent without any intrinsic motivation and motivation crowding out. Setting $\gamma = 0$ with $k > 0$ reduces to model to the one used by Besley and Ghatak (2005) where they assume that motivated agents differ from the standard agent only in terms of obtaining a higher marginal benefit of effort. This case with no motivation crowding out but with intrinsic motivation is illustrated in blue. The case with both intrinsic motivation and motivation crowding out is represented in green. This line will always lie below the blue line but may lie above or below the red one depending on the strength of motivation crowding out effect relative to intrinsic motivation, that is, depending on if k is greater or less than $\gamma\beta g$. The case illustrated here is for $k > \gamma\beta g$.

[17] In the data, I only observe if the effort of a teacher increased or remained unchanged.

[18] I assume that the teacher experiences motivation crowding out only if he/she increases effort and this is zero otherwise. Hence the motivation crowding out effect $\gamma\beta g\tilde{e}$ is always non-negative.

[19] Mathematically, $\frac{du_T}{d\beta} < 0$ for $\gamma > 1$.

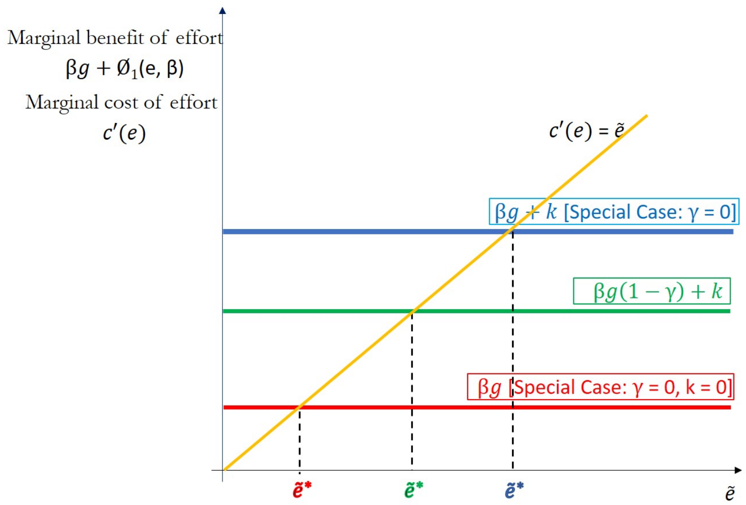


Figure 1: Marginal benefit of effort under different parameter restrictions

Before deriving the optimal incentive intensity, it is helpful to consider the effort in the first-best scenario where effort is observable and contractible.

Benchmark The first-best effort $\tilde{e}^{**}$, derived by maximising the joint surplus of the principal and the agent $fe + ke - \frac{1}{2}\tilde{e}^2$, is given by $\tilde{e}^{**} = f + k$. This gives the joint social surplus of $\frac{1}{2}(f+k)^2 + (f+k)e_0$.[20]

The following proposition states the properties of the optimal incentive intensity for a motivated agent when effort is unobservable and compares it to that for a self-interested agent. All formal proofs are provided in the appendix for the reader's perusal.

Proposition 1. *(i) The optimal incentive intensity is decreasing in a teacher's intrinsic motivation k in the presence of a motivation crowding out effect and it is also decreasing in the motivation crowding out parameter γ. (ii) Further, the extent to which the optimal incentive intensity decreases in one's intrinsic motivation is greater in the extent of the motivation crowding out effect and vice-versa. (iii) Optimal contracts consist of a lower β for a motivated agent than a self-interested agent.*

The fact that the optimal incentive intensity is lower for the motivated agent than the self-interested agent reflects two effects: (1) as the motivated agent has a higher marginal benefit of effort for any given β due to intrinsic motivation, the principal needs to provide a lower incentive to this type of agent to incentivise a given level of effort and (2) since a higher β induces a higher motivation crowding out effect $\gamma\beta g\tilde{e}$, it is optimal for the principal to set a β which takes into account this trade-off between incentives and motivation crowding out.

This tradeoff also means that the optimal incentive intensity is decreasing in k in the presence of motivation crowding out since a higher intrinsic motivation level k dictates a higher

[20]This uses $\tilde{e} = e - e_0$ to reformulate the problem of choosing e as that of choosing $\tilde{e}$. Since $\tilde{e} = e - e_0$ where e_0 is predetermined at the time of choosing an effort level e, choosing e and choosing $\tilde{e}$ are equivalent.

effort choice by incentive compatibility, which in turn means that, ceteris paribus, there is a higher motivation crowding out effect $\gamma\beta g\tilde{e}$ through greater effort. This would lower the principal's payoff as α must be increased in order to satisfy the participation constraint and the principal will therefore lower β optimally to alleviate this tradeoff.[21] Hence, in the presence of motivation crowding out, intrinsic motivation represented through k and extrinsic motivation through β act as substitutes for the principal in the optimal incentive contract.[22] Optimal intensity is also decreasing in γ since a higher γ induces a higher motivation crowding out effect all else equal and the optimal β set therefore reflects this incentives-motivation crowding out tradeoff.

Corollary 2. *The optimal contract in the presence of motivation crowding out with effort being unobservable, achieves an effort level which is lower than the first-best case.*

This distortion away from the first-best scenario results due to two effects when effort is unobservable. First, with motivation crowding out, the effective incentive intensity that the agent faces ($\beta(1-\gamma)$), is always lower than that which the principal sets (β). This is the direct effect which means that the effort induced will be lower, all else equal.[23] Second, as Proposition 1 states, optimal incentive intensity is decreasing in γ as the principal faces a tradeoff between incentives and motivation crowding out, which endogenously further reduces the effort induced as the optimal intensity is lowered. Hence, only the second-best effort level is achievable when effort is unobservable as the principal faces a tradeoff between incentives and motivation crowding out.

One can also note that if $\gamma \to 1$, that is, motivation crowding out effect tends to 100%, it is optimal to set $\beta^* = 0$ as the agent is not responsive to any incentives given the one-to-one motivation crowding out.

While the focus of the paper is on modelling the effects of incentives for teachers, the model outlined in this section can be used for occupations where self-motivation may play an important role. It helps to understand when monetary incentives may not be as useful if they displace the intrinsic motivation of the agent to a great extent. It may therefore be applied in a broader contextual setting to healthcare or other public sector employees with the government playing the role of the principal. The results derived here also help to illustrate how the implications of the standard model compare with a model which allows for the possibility of motivation crowding out. Continuing with this framework, I now turn to analyse the provision of incentives when these are provided based on group performance.

[21] Note that without a motivation crowding out effect, the optimal intensity is $\beta^* = f/g$, that is, it is independent of intrinsic motivation level k. This is because without motivation crowding out, the principal does not have to compensate the agent through a higher α for the reduction in utility that the agent experiences due to motivation crowding out.

[22] One can observe through the incentive compatibility constraint that β and k both incentivise the agent to exert a higher effort all else equal and therefore act as substitutes for motivating effort for the agent but the proposition goes further and derives that in the presence of motivation crowding out, they will act as substitutes for the principal under the optimal incentive contract.

[23] This is analogous to the deadweight welfare loss that results when using distortionary taxation. In the presence of motivation crowding out, inducing effort by providing monetary incentives cannot achieve the first best.

3.2 Group Incentives

Compared to the extensive literature on individual incentives for agents, there is relatively sparse theoretical and experimental literature on team incentives. This section lays out a simple framework for analysing the provision of team incentives for teachers. It builds on the model used in Bandiera et al. (2013) for analysing team incentives in the fruit-picking industry by allowing for intrinsic motivation and motivation crowding out. Unlike Bandiera et al. (2013), the framework outlined here takes the team composition as given, and focuses instead on the effects of team incentives on the effort of an agent.

Consider now the scenario where teachers are incentivised based on the average increase in test scores of students in their team. Let the average increase in the test scores in a team be denoted by $\overline{S_t} = \frac{\Sigma_i \Sigma_j \tilde{S}_{ijt}}{\Sigma_i m_i}$ where j indexes the students of a teacher (indexed by i) and m_i denotes the total number of students of a teacher. Here, $\tilde{S}_{ijt} = S_{ijt} - S_{ijt-1}$ denotes the change in a student's test score. To keep the analysis simple, I abstract from heterogeneity in the number of students a teacher teaches in a team. Hence the expected average increase in the test scores for a team becomes $E[\overline{S}|e_i, e_{-i}] = \frac{\Sigma_i g_i \tilde{e}_i}{n}$ where n denotes the number of teachers in the team. Hence, the measure based on which teachers are incentivised is a weighted average of their efforts, with the weights given by their marginal product of effort.[24]

The principal cares about the total effort of teachers in a team, treating individual teacher efforts as substitutes, that is, the principal's overall expected payoff will equal $u_G = f(\Sigma_i e_i) - \Sigma_i \alpha_i - n\beta\overline{S}$. For analytical tractability, I assume that a team comprises of two teachers, that is, $n = 2$.[25]

When providing group incentives it is important to consider the extent to which the teachers can internalise the externality that their effort has on the other team members' earnings or weight the other team members' earnings from the incentive scheme. I assume that a teacher acts independently based on his/her own intrinsic preferences when provided with group incentives but I allow for the teacher's effort choice to be affected by the extent to which he/she internalises the externality of her effort on the other teachers' incentive payment in the sense made precise below.

A teacher's utility is $u_{T_i} = \alpha_i + \beta\overline{S} + \phi(e, \beta) - c(e)$. The teacher however may choose her effort taking into account the effect it has on the other teacher's incentive payments. In other words, a teacher can to some extent internalise the externality that his/her effort choice has on the other teacher's earnings. The teacher therefore chooses effort to maximise $\alpha_i + \beta(1 + \tau_i(n-1))\overline{S} + \phi(e, \beta) - c(e)$.[26]

---

[24] As before, this assumes that the idiosyncratic errors in the technology of student test scores are mean zero and the principal is risk neutral.

[25] In Andhra Pradesh, a typical primary school has three teachers (Muralidharan and Sundararaman, 2011).

[26] Note that this parametrization ensures that whilst the teacher experiences motivation crowding out from the provision of monetary incentives, he/she does not experience motivation crowding out from the provision of monetary incentives to the other team member. Mathematically, this is equivalent to not having e_{-i} as an argument in the motivation function.

Here $\tau_i \in [0,1]$ is a scale parameter which is the relative weight a teacher puts on each of the other team members' gain from the team's performance, that is, it captures the extent to which a teacher internalises the effect of her performance on the other team member's earnings from this performance pay scheme.[27][28] This parameter can be proxied by the strength of cooperation amongst the teachers in a team.[29] Here $\tau_i = 0$ denotes no cooperation as the teacher only cares about his/her own earnings from the performance pay scheme and $\tau_i = 1$ denotes full cooperation to the extent that the teacher assigns equal weight to each of the other team members' earnings as much as his/her own when choosing effort.[30][31]

As before, $c(e) = \frac{1}{2}\tilde{e}_i^2$.[32] Let the motivation function be $\phi(e,\beta) = k_i e_i - \gamma\frac{\beta}{2}g_i\tilde{e}_i$.[33] The objective function that the teacher maximises, assuming there are two teachers in a team ($n = 2$), can thus be written more explicitly as[34]

$$\alpha_i + \beta(1+\tau_i)\left(\frac{g_i\tilde{e}_i + g_{-i}\tilde{e}_{-i}}{2}\right) - \gamma_i\frac{\beta}{2}g_i\tilde{e}_i + k_i e_i - \frac{1}{2}\tilde{e}_i^2$$

The timing of events is as follows.

1. The government stipulates the contract including the performance related pay com-

[27]The parameter τ can also be interpreted here as the extent to which the individual teacher's objectives are aligned with that of the team or the extent to which the teacher is altruistic towards the incentive pay earnings of the other team member. Mathematically, this parameter acts as a reduced form to capture the extent to which the teacher chooses effort to maximise joint utility, which in case of $n = 2$, is $u_{T_i} + \tau_i u_{T_{-i}}$.

[28]Since there is no ex-ante reason to suppose that the dis-utility the teacher experiences due to explicit performance pay via motivation crowding out differs systematically with a teacher's altruism, the parameter τ_i does not affect the motivation crowding out effect in the teacher's objective function. In other words, the distaste of the teacher towards explicit performance pay is only affected by the teacher's effort and the incentive intensity, and is unaffected by the teacher's altruism.

[29]As the teachers may differ in the extent to which they find the other teachers in their team cooperative, I allow this to be another dimension of heterogeneity.

[30]I assume that the extent to which a teacher internalises the effect of his/her effort on the other teacher's earnings is monotonically increasing in the extent to which he/she views (and reports) the teachers in their team to be cooperative. I therefore use the phrases 'teacher's reported group cooperation level' and 'the extent to which the teacher internalises the effect of his/her effort on the other teacher's earnings' interchangeably henceforth.

[31]Note also that if $\tau_i = 1$ for all the teachers in a team, this represents the case where all the teachers maximise the team's utility, that is, it would achieve the solution that could be achieved if all teachers acted cooperatively to achieve the full efficiency solution under incentive provision.

[32]Note that since there is no technological dependence between the agents and they are assumed to only act based on their internalisation of the externality, proxied by their reported group cooperation level and not necessarily their actual cooperation, I assume that their cost of effort remains unchanged. In other words, τ_i does not affect $c(e)$.

[33]Since the teacher's effective incentive intensity is halved as the incentive payment is based on group performance, the motivation crowding out effect is also halved compared to individual-level incentives case where this was $\gamma\beta g_i\tilde{e}_i$. As before, I assume that the teacher only experiences motivation crowding out if he/she increases effort, that is, $\gamma\frac{\beta}{2}g_i\tilde{e}_i$ is always non-negative.

[34]Mathematically, the provision of group incentives is modelled here as affecting the incentive compatibility constraint for the teachers but not the participation constraint.

ponent based on the improvement of the average student test score of both teachers in a team.

2. A teacher individually accepts or rejects the contract. If a teacher rejects the contract, both teachers get the reservation utility $\overline{u}$ from an outside option. If the contract is accepted by both teachers, a teacher acts independently (non-cooperatively) and maximises his/her expected utility choosing their privately-observed effort subject to the participation constraint $E[u_{T_i}|e_i, e_{-i}] \geqslant \overline{u}$.

3. Random shocks (ε, η) are realised but unobserved by the government. Both the government and the teacher can observe the student test scores and teachers are paid in accordance with the stipulated contract terms.

In addition to the three dimensions of heterogeneity amongst teachers - their marginal product of effort on raising test scores or 'ability' g, the extent to which they experience motivation crowding out γ, and their intrinsic motivation k - teachers also differ in terms of the extent of internalising the effect of their effort on the other team member's earnings τ, which is proxied here by their reported group cooperation level. As before, I assume that teachers know their intrinsic preferences but that they do not know the other teacher's intrinsic preferences or the extent to which the other teacher internalises the externality.

Whilst the provision of incentives at the level of the team reduces the private marginal benefit of effort for a teacher as part of the benefit ensues to the other team member, this effect may be mitigated by the extent to which the teacher cares about their joint earnings as the following lemma shows.

Lemma 2. *Incentive compatibility for the teacher dictates*[35]*:*

$$\tilde{e}_i = \frac{(1 - \gamma_i + \tau_i)\beta g_i}{2} + k_i$$

There are a number of interesting points to note here. Firstly, if there is no internalisation of the externality among the teachers in team, that is $\tau_i = 0$, then the effort level of a teacher will be lower than that of a teacher who is given incentives based on his/her own performance holding all else equal. This is the standard prediction derived from the fact that the effective incentive intensity for the teacher is weaker when based on a team output. In this case, providing individual incentives would have been more effective to raise effort.

Secondly and more interestingly, if $\tau_i = 1$, denoting that there is full internalisation of the externality, then the effort level of a teacher can be higher following the provision of team incentives compared to the provision of individual-level incentives. This is because of two reasons: (1) the extent of the motivation crowding out effect is now weaker given that the effective incentive intensity for the agent is halved and (2) the teacher obtains utility from his/her team member's earnings on which he/she does not experience any motivation crowding out. More generally, if $\tau_i > 1 - \gamma_i$, then providing group incentives

[35]The concavity of the objective function ensures that the first order condition is sufficient for a maximum.

can dominate providing individual-level incentives if the objective is to increase effort. We can therefore observe the following result.

Corollary 3. *For a given β, providing group incentives could be better than individual-level incentives for incentivising effort for teachers who have a high level of motivation crowding out and internalise the effect of their effort on the other teacher's earnings to a large extent.*[36]

The model therefore helps to shed light on why, contrary to the predictions of the theoretical models which abstract from motivation crowding out and the extent to workers can internalise the effect of their effort choice on the other team member's earnings, provision of team based incentives could improve upon individual incentives if the objective of the principal is to increase effort. Hence, the extent to which the provision of incentives can affect a teacher's effort levels is contingent on the form that these are provided in - as individual or group incentives - and the extent to which teachers internalise the externality of their effort on the other team member.

Corollary 4. *The responsiveness of a teacher's effort to an increase in the incentive intensity at the optimum is greater for a teacher with a higher level of internalisation of the externality all else equal.*

This follows from noticing that $de_i/d\beta$ is increasing in τ. Before deriving the optimal group incentive intensity for a team, it is helpful to first consider the benchmark case of observable and contractible effort.

Benchmark. The optimal level of effort for each team member if effort is observable and contractible is the same as in the case where the principal contracts individually with a teacher. This can be derived by maximising the joint social surplus

$$f(\tilde{e}_1 + \tilde{e}_2) + k_1\tilde{e}_1 + k_2\tilde{e}_2 - \frac{1}{2}(\tilde{e}_1^2 + \tilde{e}_2^2) + f(e_{10} + e_{20}) + k_1 e_{10} + k_2 e_{20}$$

with respect to each agent's effort. This gives the optimal effort level $\tilde{e}_i^{**} = f + k_i$ as in the individual-level incentives case. The resultant joint social surplus here is $\frac{1}{2}(f + k_1)^2 + \frac{1}{2}(f + k_2)^2 + f(e_{10} + e_{20}) + k_1 e_{10} + k_2 e_{20}$.

The following proposition characterises the properties of the optimal incentive intensity for group-level incentives.

Proposition 2. *(i) The optimal incentive intensity is decreasing in intrinsic motivation k_i of a teacher in the presence of motivation crowding out. (ii) The ordering of optimal incentive intensity in four special cases depending on the average level of motivation crowding out and internalisation of the externality in the team $\{\overline{\gamma}, \overline{\tau}\}$ is as follows:*

$$\beta^*_{\overline{\gamma}=0,\overline{\tau}=0} > \beta^*_{\overline{\gamma}=0,\overline{\tau}=1} > \beta^*_{\overline{\gamma}=1,\overline{\tau}=1} \geqslant \beta^*_{\overline{\gamma}=1,\overline{\tau}=0}$$

[36] The precise condition for this, as stated above, is $\tau_i > 1 - \gamma_i$.

with the last inequality being strict if the optimal incentive intensity is non-zero in the third case. The corresponding ordering of the teacher effort level is:

$$\tilde{e}^*_{i,\{\overline{\gamma}=0,\overline{\tau}=0\}} = \tilde{e}^*_{i,\{\overline{\gamma}=0,\overline{\tau}=1\}} > \tilde{e}^*_{i,\{\overline{\gamma}=1,\overline{\tau}=1\}} \geqslant \tilde{e}^*_{i,\{\overline{\gamma}=1,\overline{\tau}=0\}}$$

with the last inequality being strict if β^* is non-zero in the third case.(iii) If the teachers in a team are identical in terms of their ability and extent to which they experience motivation crowding out, then the optimal incentive intensity is decreasing in the extent to which they experience motivation crowding out.

The results show that as in the individual incentives case, the optimal incentive intensity is decreasing in both the teachers' intrinsic motivation if they experience some motivation crowding out. The intuition behind the result is as under the individual-level incentive scheme. A higher intrinsic motivation dictates a higher effort level for a teacher, which in turn means a higher motivation crowding out effect for that teacher of $\beta\gamma_i g_i \tilde{e}_i/2$. Since the principal will have to increase α_i to compensate for this, the principal optimally takes this into account when setting the incentive intensity. Thus, as before, in the presence of motivation crowding out, intrinsic motivation through k_i and extrinsic motivation through β act as substitutes for the principal in the optimal incentive contract.

The overall effect of a teacher's level of internalisation of the externality τ_i is theoretically ambiguous as the indirect effects of this on the principal's net payoff are conflicting.[37] Whilst a higher τ_i raises principal's payoff by inducing a higher effort level by the agent and increasing β would increase this effect further, it also indirectly raises the motivation crowding out effect and the disutility of effort experienced as both of these are increasing in effort and increasing β would also increase these effects further. Given these conflicting indirect effects on the principal's net payoff of raising β when τ_i increases, and further, as the optimal group incentive intensity also has to take into account the effect of changing the other teacher's effort incentives, it is not possible to derive general analytical statements regarding the comparative statics of the optimal incentive intensity with respect to this parameter.

The overall effect of an increase in the teacher's motivation crowding out level is also theoretically ambiguous. Although the direct effect of a higher motivation crowding out on the principal's net payoff is negative and the principal has to raise α_i to meet the agent's participation constraint all else equal, under the group incentive scheme the principal also has to balance the effect of lowering the β in response to this by considering the effect this would have on the other teacher's effort incentives.

Therefore, as the optimal incentive intensity takes the form of a weighted average of both teachers' abilities and intrinsic motivation - with the weights depending on the teachers' motivation crowding out and the teacher's extent of internalisation of the externality as derived in the proof above - this precludes any general analytical statements regarding the comparative statics with respect to a teacher's extent of internalisation of the externality and motivation crowding out. Furthermore, this also illustrates the difficulty in setting

[37] Note that the parameter only affects the agent's and principal's payoff through the effort level (that is, indirectly) as it enters the incentive compatibility constraint for the agent but not the participation constraint.

the optimal group incentive intensity as it highlights the demanding informational requirements for the principal and the difficulty of tailoring the optimal incentive intensity in response to an individual teacher's preferences.

Using simulations I find that teachers have to differ substantially along the three dimensions of heterogeneity $\{\gamma, g, k\}$ for the incentive intensity to be increasing in a teacher's motivation crowding out.[38] As teachers in a school are unlikely to differ substantially along all of these dimensions of heterogeneity, it is perhaps reasonable to conclude that the incentive intensity is decreasing in a teacher's motivation crowding out more generally. Further, the simulations also suggest that the derivative of the optimal incentive intensity with respect to the extent to which the teacher internalises the externality depends on the teacher's motivation crowding out parameter; more precisely, the cross derivative $\frac{\partial^2 \beta^*}{\partial \gamma_i \partial \tau_i} > 0$.[39]

---

[38]The necessary condition for the incentive intensity to be decreasing in a teacher's motivation crowding out γ_i is

$$-\frac{fg_i^2(1+\tau_i-\gamma_i)^2}{2} - \frac{fg_{-i}(1+\tau_{-i}-\gamma_{-i})}{2}[g_{-i}(1+\tau_{-i}+\gamma_{-i}) - 2\gamma_i g_i] \\ - k_i g_i^2 \frac{(1+\tau_i)^2-\gamma_i^2}{2} - k_i g_{-i}\frac{(1+\tau_{-i})^2-\gamma_{-i}^2}{2} - \gamma_i^2 g_i^2 k_i - \gamma_i\gamma_{-i}g_i g_{-i}k_{-i} \leqslant 0$$

Using 10,000 independent simulations of $\gamma \sim U[0,1]$, $g \sim U[0,9]$, $k \sim U[0,5]$ and $\tau \sim U[0,1]$ for each teacher along with uniformly drawn values of $f \sim U[0,7]$, I find that the chance of the optimal incentive intensity β^* being increasing in either teacher's motivation crowding out intensity is close to zero (38 cases in 10,000 simulations). Moreover, using these simulations, I also find that even if the optimal incentive intensity is increasing in a teacher's motivation crowding out, it is decreasing in the other teacher's motivation crowding out.

Furthermore, for the few cases where the incentive intensity is increasing in a teacher's motivation crowding out, the teacher's motivation crowding out and ability, that is the product $\gamma_i g_i$ are both much higher than that of the other teacher's and the teacher's intrinsic motivation k_i is much lower compared to that of the other teacher. Hence, the teachers have to be considerably different in these three dimensions of heterogeneity $\{\gamma, g, k\}$ simultaneously in order for the unlikely case of the incentive intensity being increasing in one teacher's motivation crowding out to materialise. Given the almost negligible chance of encountering this as teachers in a school are likely to be similar along one or more of these dimensions of heterogeneity, it is perhaps reasonable to conclude that the incentive intensity is decreasing in a teacher's motivation crowding out more broadly.

[39]The necessary condition for the optimal intensity to be increasing in τ_i is

$$-\frac{fg_i^2(1+\tau_i-\gamma_i)^2}{2} + \frac{fg_{-i}(1+\tau_{-i}-\gamma_{-i})}{2}[g_{-i}(1+\tau_{-i}+\gamma_{-i}) - 2g_i(1+\tau_i] \\ + k_i g_i^2 \gamma_i(1+\tau_i) + k_{-i}g_i g_{-i}\gamma_{-i}(1+\tau_i) \geqslant 0$$

Using 10,000 independent simulations of $\gamma \sim U[0,1]$, $g \sim U[0,9]$, $k \sim U[0,5]$ and $\tau \sim U[0,1]$ for each teacher and $f \sim U[0,7]$, I find that the chance of the optimal incentive intensity β^* (when it is non-zero) being decreasing in a teacher's level of internalisation of the externality is around a third (3403 cases in 10,000 simulations). Moreover, the chance that it is decreasing in both the teachers' level of internalisation of the externality is even lower at 0.15. Furthermore, as the teacher's motivation crowding out tends to one, the chance that optimal incentive intensity is decreasing in τ_i, when it is non-zero, tends to zero (0.04). This suggests that the derivative of the optimal incentive intensity with respect to the level of internalisation of the externality of a teacher is likely to be positive and increasing in his/her motivation crowding out, that is, the cross derivative $\frac{\partial^2 \beta}{\partial \gamma_i \partial \tau_i} > 0$.

The following corollary helps to illuminate how the effort under group incentives compares to the benchmark case of observable and contractible effort.

Corollary 5. *Under the best possible scenario facing the principal when providing group incentives, that is, when teachers face no motivation crowding out ($\overline{\gamma} = 0$) and internalise the effect of their efforts on each other's earnings fully ($\overline{\tau} = 1$)*[40]*, first best level of effort is achieved for both teachers if and only if both teachers are identical in terms of their ability. Otherwise, the sum of the efforts achieved is less than the first best level.*

This result shows that unless the teachers in a team are identical in terms of their ability, even under the best possible scenario facing the principal when providing group incentives, the sum of the effort levels achieved under the optimal contract is strictly less than that in the first-best case. Further, as the effort level of a teacher is decreasing in the motivation crowding out of a teacher (through incentive compatibility) and since the simulations show that the optimal incentive intensity is also almost always endogenously decreasing in the motivation crowding out parameter of a teacher, the sum of the efforts of the teachers under the group incentive scheme will be strictly less than that of the first best case. Hence this shows that if teachers experience motivation crowding out, the sum of effort levels achieved in a team when providing group incentives is likely to be lower than the first best case more generally.

Moreover, in this best possible scenario when teachers do not experience motivation crowding out and internalise the effect of their effort on the other team member's earnings fully, providing individual-level incentives would have instead achieved the first best effort levels for both the teachers and therefore would be the optimal incentive design. This holds true because of the inability to tailor the incentive intensity for each of the group members; instead, as the proof for Proposition 2 illustrates, the optimal incentive intensity takes a weighted average of both teachers' abilities and intrinsic motivation with the weights depending on the teachers' motivation crowding out and extent of internalisation of the externality. I now turn to explicating how the theoretical framework outlined here can be used to analyse the experimental results of teacher performance pay.

4 Estimation of the structural model

4.1 Data and descriptive statistics

Teachers in both incentive and control group schools were interviewed about whether they gave any special preparation to their students during the year. These interviews can help shed light on the changes in their effort levels when they were incentivised based on student test scores. In particular, each year teachers were asked if they (1) assigned extra homework; (2) assigned extra class work; (3) gave extra exam practice; (4) gave extra classes or taught beyond school hours; and (5) paid special attention on weaker

[40] This is the best possible scenario for the principal when providing group incentives in the sense that all else equal, the principal needs a lower incentive payment to incentivise a given level of effort.

students.[41][42] These are coded as binary variables in the dataset. In order to measure the change in effort, I aggregate these variables to generate a six-point support for the choice of change in effort, i.e. $\tilde{e} \in \{0, 1, 2, 3, 4, 5\}$.[43] As I have argued earlier, since these activities are likely to be complementary in raising test scores, I take effort to be uni-dimensional in my analysis here.[44]

The figure below plots the relative frequency distribution of effort and change in average student test score in all the schools.[45] The median effort choice was higher at 2 in the treatment schools than the control schools, where the median was 1. The figure below also shows that the relative frequency of effort choice of zero was comparatively much lower in the treatment group than the control group.

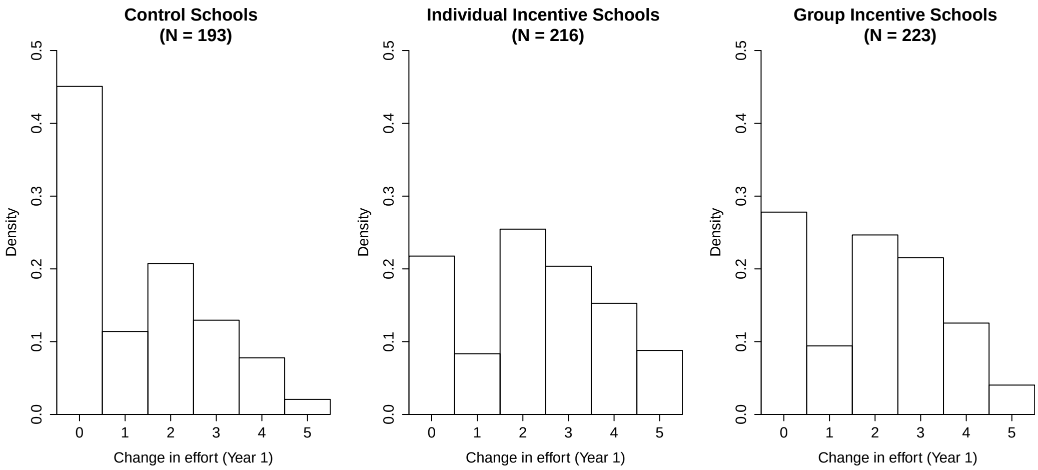


Figure 2: Changes in teacher effort. N refers to the number of teachers.

The table below provides the ordinary least squares intent to treat estimates for effort.[46] Using control group schools as the base category, the dummy variables 'Individual-

[41]Based on the teaching activities recorded using unannounced visits in treatment and control schools, the authors report that there were no substantial and significant differences in measures of classroom activities (Muralidharan and Sundararaman, 2011). Hence I focus on measuring effort using these self-reported activities.

[42]Note that these variables record the *increase* in effort, that is, the change in effort levels with respect to the teacher's effort last year.

[43]For ease of description, I shall refer to this index as the effort choice of the teacher henceforth in this section.

[44]Since the responses are self-reported, any measurement error in reporting is likely to be systematically linked across these 5 questions. Hence, I avoid using principal component analysis based on these 5 questions to obtain a measure of effort and simply aggregate the responses to form an index.

[45]Data on student performance and teacher interviews was first matched, linking students to their teachers. Student observations for which improvement in test scores were missing were dropped from the analysis given that these would not enter into calculations of the performance of a teacher. Analysis presented in this section is based on these combined datasets for control and treatment schools. For computing the average student test score, I average the data over maths and language test scores of a teacher's students if the teacher teaches both subjects.

[46]As I treat effort as being unidimensional in this setting but only observe discrete values of the

incentives' and 'Group incentives' pick up the differences in mean effort choices. Results show that the difference in effort between the the control group and the two treatment groups is statistically significant.

Table 1: Intent to Treat Estimates

	Effort ($\tilde{e}$)
Individual-incentives	0.874***
	(0.212)
Group incentives	0.676**
	(0.208)
Constant	1.417***
	(0.156)
Observations	475
R^2	0.301

Mandal (subdistrict) fixed effects included.
Standard errors (in parentheses) are clustered at the school level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A chi-squared test of difference in distributions of effort choices between the control and group incentive schools gives a test statistic of 17.61 with a p-value of 0.003. The chi-squared statistic for effort choice distribution between individual-level incentive and control group schools is 35.29 with a p-value of 0.000. This again provides evidence that the teachers in the two treatment groups made effort choices generating a distribution that is statistically different from that in the control schools. The corresponding chi-squared test for difference in distribution of effort choice between individual-level and group incentive schools, however, is not statistically significant with a test statistic of 6.34 and a p-value of 0.275.

The mean improvement in test scores was higher in treatment schools than control schools. This was 15 percentage points in individual incentive schools and 13 percentage points in group incentive schools whilst the corresponding statistic was 10 percentage points in control group schools. The median was also two percentage points higher in group incentive schools and three percentage points higher in individual incentive schools compared to the control group.[47] This provides suggestive evidence supporting the authors' hypothesis that incentive schools performed better as teachers increased their effort intensity in a way that was not captured through classroom observation.

4.2 Set Identification of Teacher Heterogeneity

I now turn to identify the structural parameters in the theoretical model using this data. Different sets of structural parameters are identifiable for the control and treatment groups given the assumptions about what factors affect the choices made by the teachers in each

dependent variable, I interpret these estimates qualitatively.

[47] Section 7.3 in the Appendix plots the distribution of test score improvement across the treatment and control schools.

scenario.[48] As the effort choice in this setting is observed as a discrete rather than a continuous variable, instead of using the incentive compatibility constraint which derives the effort choice of the teacher assuming continuity of effort, I work directly with the utility function for structural estimation.[49]

I depart from the standard approach of treating unobserved heterogeneity as a multivariate error term as is often motivated in a random utility framework; instead I use a set identification strategy for the following reasons. If we denote $\theta_{\mathbf{i}}$ as the vector of heterogeneous parameters and $\mathbf{x_{ij}}$ as the observable characteristics driving the effort choice j of the teacher i, we can write the utility function of the teacher as specified in the theoretical section as $U(\theta_{\mathbf{i}}, \mathbf{x_{ij}})$. It is therefore important to note firstly that the theoretical framework does not specify any additional structure such as the joint distribution of $\theta_{\mathbf{i}}$. Standard estimation techniques, such as Maximum (Simulated) Likelihood, rely on parametric distributional assumptions regarding the joint or marginal distribution of the unobservable heterogeneity and therefore necessitate additional untestable assumptions regarding the nature of underlying unobservable heterogeneity (Tamer, 2010). Instead, I wish to examine what may be learned from the data solely based on the theoretical framework which only provides us with inequalities relating a teacher's preferences and effort choice here. In other words, my objective here is to derive the necessary conditions that the unobservable heterogeneous parameters must satisfy if the data were to be generated by the theoretical framework outlined.[50][51]

Secondly, multiple dimensions of unobservable heterogeneity necessarily mean that a choice can be rationalised by multiple combinations of the heterogeneity parameters. For example, a teacher may have chosen a low effort level either because he/she has a low level of intrinsic motivation and experiences a low level of motivation crowding out, or because he/she has a high level of intrinsic motivation but experiences a very high level of motivation crowding out. In other words, different combinations of the heterogeneous parameters are observationally equivalent (Lewbel (forthcoming)). In light of the

[48] I continue to use a static framework to interpret the effort choices of teachers. In the experimental setting, teachers were informed that the experiment is likely to, contingent on government funding, but not guaranteed to roll on for the next year. Since I do not observe teachers' expectations regarding the continuation of the program, the identified sets here will be the relevant ones for at least those teachers who expected the program to last for one year or who placed a high probability on this and/or for those teachers who might have a very low discount factor.

[49] For simplicity, I have ignored the threshold of 5% that was set in the first year. This is without loss of generality because the threshold was set to account for the fact that students were first tested following the summer holidays and would have been unlikely to have the academic material fresh in their minds. It was therefore expected that student test scores would increase by 5% from baseline during the academic year even without the teacher's input. This means that effectively the teachers faced a linear payment schedule for their effort and we can therefore abstract from the literal piece-wise linear bonus scheme set in the first year when analysing teachers' effort choice. Further, since the bonus was not discontinuous at the threshold, teachers would have been unlikely to try to game the scheme. This is confirmed empirically as the authors report that the distribution of scores at the end of the first year of the program did not have a gap below the threshold of 5%, which would have otherwise been indicative of gaming behaviour.

[50] I therefore treat the heterogeneous parameters simply as random variables and derive bounds that this set must satisfy given a teacher's effort choice.

[51] This approach is often used in the revealed preference empirical literature which uses systems of inequalities based on Afriat's theorem to test for utility maximisation without relying on any structural functional form assumptions.

recent literature emphasising the importance of understanding the assumptions driving identification in models and acknowledging the limitations of what the data can allow us to identify (Tamer, 2010, Keane, 2010, Lewbel (forthcoming)), I therefore adopt a set or partial identification strategy whereby I derive bounds on the sets of heterogeneity consistent with each effort choice.[52] Set identification is proved by construction in what follows.

4.2.1 Control Group

Let $U_{ij} = \alpha_i + k_i\tilde{e}_{ij} + k_i e_{i0} - \frac{1}{2}\tilde{e}_{ij}^2$ be the expected utility teacher i derives from effort choice of $j \in \{0, 1, 2, 3, 4, 5\}$.[53]Since the utility functions are unobserved, but the choice of effort is, I use a a latent variable framework for a discrete ordered choice to identify the set of structural parameters consistent with each effort choice.[54] We will observe an effort choice $\tilde{e}^*_{j+1}$ if $U_{ij+1} \geqslant U_{ij}$, that is, if for example we observe effort choice of 1, it must be that $U_{i1} \geqslant U_{i0}$ and $U_{i1} \leqslant U_{i2}$. We can therefore obtain the set of values of k consistent with each effort choice for the control group by rearranging these inequalities.

This gives:

$$
\begin{aligned}
k_i \leqslant \frac{1}{2} \quad &\Leftrightarrow \quad \tilde{e}^* = 0 \\
\frac{1}{2} \leqslant k_i \leqslant \frac{3}{2} \quad &\Leftrightarrow \quad \tilde{e}^* = 1 \\
\frac{3}{2} \leqslant k_i \leqslant \frac{5}{2} \quad &\Leftrightarrow \quad \tilde{e}^* = 2 \\
\frac{5}{2} \leqslant k_i \leqslant \frac{7}{2} \quad &\Leftrightarrow \quad \tilde{e}^* = 3 \\
\frac{7}{2} \leqslant k_i \leqslant \frac{9}{2} \quad &\Leftrightarrow \quad \tilde{e}^* = 4 \\
\frac{9}{2} \leqslant k_i \quad &\Leftrightarrow \quad \tilde{e}^* = 5
\end{aligned}
$$

It is possible to identify the distribution of intrinsic motivation in this setting since in the absence of incentives, parameters β, g, γ and τ do not enter into the utility function. Since the utility function is invariant to a monotonic transformation, α and k are divided by 1000 to make comparison with the parameter set derived for teacher choice in the treatment schools possible. Matching this set with the effort choice distribution that we observe

[52]As expected, given observational equivalence of different combinations of heterogeneous parameters, some of which enter non-separably in the utility function, point identifying using a Maximum Simulated Likelihood strategy does not work successfully as the convergence is not global and depends on the initial guess of the starting parameter values. The results for this exercise are available upon request for the reader's perusal.

[53]The term $k_i e_{i0}$ appears here as the term $k_i e_{ij}$ is replaced by $k_i\tilde{e}_{ij}$.

[54]One can note that since the effort choice of the teacher will be invariant to any monotonic transformations of the utility function, using a linear utility is not imposing a functional form restriction on the utility function.

in control school, we get a distribution for intrinsic motivation of a teacher in control group schools. The diagram below illustrates this and suggests that the distribution of intrinsic motivation among teachers is bimodal with teachers having either very low or medium levels of intrinsic motivation. Given randomization, the distribution of intrinsic motivation in the treatment schools is likely to be the same.

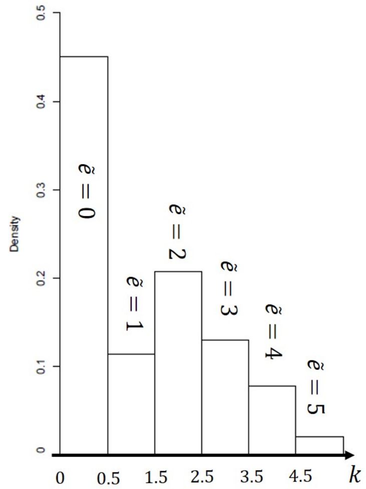


Figure 3: Density of intrinsic motivation. This is derived by combining the distribution of effort choice in control school from Figure 2 with the derived parameter set of intrinsic motivation consistent with each effort choice.

4.2.2 Individual-level Incentives

Let $U_{ij} = \alpha_i + \beta(1-\gamma_i)g_i\tilde{e}_{ij} + k_i\tilde{e}_{ij} + k_i e_{i0} - \frac{1}{2}\tilde{e}_{ij}^2$ be the utility teacher i derives from effort choice of $j \in \{0,1,2,3,4,5\}$.[55] Using similar reasoning as above, the three-dimensional parameter set of teacher heterogeneity that is consistent with the observed effort choice can be derived as:

$$\begin{aligned}
\beta(1-\gamma_i)g_i + k_i \leqslant \frac{1}{2} \quad &\Leftrightarrow \quad \tilde{e}^* = 0 \\
\frac{1}{2} \leqslant \beta(1-\gamma_i)g_i + k_i \leqslant \frac{3}{2} \quad &\Leftrightarrow \quad \tilde{e}^* = 1 \\
\frac{3}{2} \leqslant \beta(1-\gamma_i)g_i + k_i \leqslant \frac{5}{2} \quad &\Leftrightarrow \quad \tilde{e}^* = 2 \\
\frac{5}{2} \leqslant \beta(1-\gamma_i)g_i + k_i \leqslant \frac{7}{2} \quad &\Leftrightarrow \quad \tilde{e}^* = 3 \\
\frac{7}{2} \leqslant \beta(1-\gamma_i)g_i + k_i \leqslant \frac{9}{2} \quad &\Leftrightarrow \quad \tilde{e}^* = 4 \\
\frac{9}{2} \leqslant \beta(1-\gamma_i)g_i + k_i \quad &\Leftrightarrow \quad \tilde{e}^* = 5
\end{aligned}$$

[55] The term $k_i e_{i0}$ appears here as the term $k_i e_{ij}$ is replaced by $k_i\tilde{e}_{ij}$.

Since the utility function is invariant to monotonic transformations, I set $\beta = 0.5$ in plotting the parameter sets below. The parameters α, which is the fixed salary, k, which represents intrinsic motivation, and the cost function are therefore to be interpreted as being scaled accordingly, that is, by a factor of 0.001. Plotting k on the z-axis and γ and g on the other axes, we can represent the parameter space and the joint parameter sets that are consistent with each observed choice as illustrated in the diagram below.

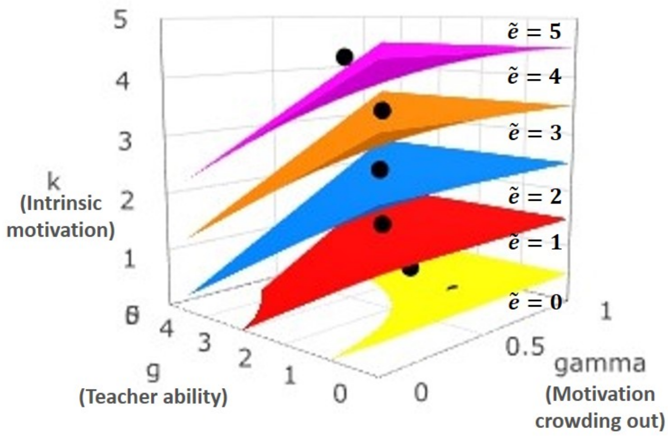


Figure 4: Sets for teacher heterogeneity in individual incentive schools (Boundaries and Barycenters). The coloured surfaces indicate the separating boundaries of the sets. Barycenters of each set $E[\gamma, g, k|e]$, marked in black dots, were calculated based on uniformly drawing 10,000 values of $\{k, \gamma, g\}$ within this cuboid and averaging over the values of the parameters which fall in each set.

The boundaries of these sets are marked in coloured surfaces in the diagram. The area between these two surfaces therefore mark the parameter set for teacher heterogeneity consistent with an observed effort choice of one. Hence, the area below the surface in yellow represents the parameter set for $\tilde{e} = 0$. The area between the red and yellow surfaces represents the parameter set for $\tilde{e} = 1$. The area between the blue and red surfaces represents the parameter set for $\tilde{e} = 2$. The area between the orange and blue surfaces represents the parameter set for $\tilde{e} = 3$. The area between the purple and orange surfaces represents the parameter set for $\tilde{e} = 4$. The area above the purple surface is consistent with choice $\tilde{e} = 5$.

Given teachers' effort choice distribution in the individual-level incentive schools, we can estimate the proportion of teachers in each of the identified parameter sets in Figure 4. The proportion of teachers in the set below the yellow surface is 0.22, between the yellow and red surface is 0.08, between the red and blue is 0.26, between blue and orange is 0.20, between orange and purple is 0.15, and above the purple surface is 0.09.

Using simulations based on uniformly drawing 10,000 values of $\{k, \gamma, g\}$ within this cuboid and averaging over the values of the parameters which fall in each set, the barycenter or centroid of these sets, $E[\gamma, g, k|e]$, are marked as black dots. These barycenters mark the center of mass of each set, that is, they plot the average values of the parameters describing teacher heterogeneity across the different sets. These can therefore be interpreted as indicative of the characteristics of the average teacher in each set. The following diagram plots these barycenters across each set consistent with each effort level.

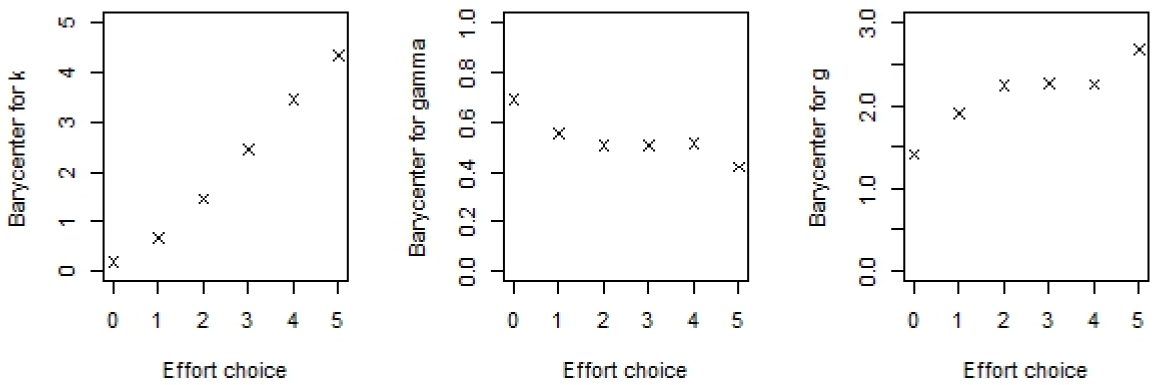


Figure 5: Barycenters for sets of teacher heterogeneity consistent with each effort choice in individual-level incentive schools. Calculations based on 10,000 simulations within the cuboid of teacher heterogeneity illustrated above.

The figure shows that intrinsic motivation of the average teacher in each of the sets is clearly increasing as we move up the sets consistent with a higher effort choice. The ability of the average teacher also exhibits a broadly increasing trend whilst the extent to which the teacher experiences motivation crowding out exhibits a decreasing trend, except for sets consistent with medium-level of effort choices where these are virtually indistinguishable, as we move up the sets consistent with higher effort choices.

4.2.3 Group Incentives

Let the utility of a teacher i from the choice of effort j be

$$U_{ij} = \alpha_i + \beta(1+\tau_i)\left(\frac{g_i\tilde{e}_{ij} + g_{-i}\tilde{e}_{-ij}}{2}\right) - \beta\gamma_i g_i\tilde{e}_{ij} + k_i\tilde{e}_{ij} + k_i e_{i0} - \frac{1}{2}\tilde{e}_{ij}^2$$

where effort choice can be $j \in \{0,1,2,3,4,5\}$. [56] Using the same reasoning as before, we can derive the three-dimensional set for teacher heterogeneity that is consistent with each effort choice conditional on the group cooperation level as follows.

$$\beta(1-\gamma_i+\tau_i)\frac{g_i}{2} + k_i \leqslant \frac{1}{2} \quad \Leftrightarrow \quad \tilde{e}^* = 0$$
$$\frac{1}{2} \leqslant \beta(1-\gamma_i+\tau_i)\frac{g_i}{2} + k_i \leqslant \frac{3}{2} \quad \Leftrightarrow \quad \tilde{e}^* = 1$$

[56] As before, the term $k_i e_{i0}$ appears here as the term $k_i e_{ij}$ is replaced by $k_i\tilde{e}_{ij}$.

$$\frac{3}{2} \leqslant \beta(1-\gamma_i+\tau_i)\frac{g_i}{2}+k_i \leqslant \frac{5}{2} \quad \Leftrightarrow \quad \tilde{e}^* = 2$$
$$\frac{5}{2} \leqslant \beta(1-\gamma_i+\tau_i)\frac{g_i}{2}+k_i \leqslant \frac{7}{2} \quad \Leftrightarrow \quad \tilde{e}^* = 3$$
$$\frac{7}{2} \leqslant \beta(1-\gamma_i+\tau_i)\frac{g_i}{2}+k_i \leqslant \frac{9}{2} \quad \Leftrightarrow \quad \tilde{e}^* = 4$$
$$\frac{9}{2} \leqslant \beta(1-\gamma_i+\tau_i)\frac{g_i}{2}+k_i \quad \Leftrightarrow \quad \tilde{e}^* = 5$$

The teacher interviews also included questions on group cooperation level of teachers. More specifically, the data collected includes three binary variables recording (1) if the teachers shared the work load; (2) if they encouraged hard work; (3) if they shared teaching methods. I aggregate these three variables to get an index between 0 and 3. Section 7.4 in the Appendix plots the histogram of this index in the group incentive schools. I then normalise this index by dividing by 3 to obtain the group cooperation level $\tau \in \{0, 1/3, 2/3, 1\}$.

The diagram below plots the two-way frequency distribution of effort choice and and reported group level cooperation, with the width of the circles being in proportion to the relative frequency of teachers. The positive slope of the line of association between effort choice and reported group cooperation level is in line with the theoretical prediction that effort choice of a teacher e_i is increasing in one's reported group cooperation level τ_i, all else equal.

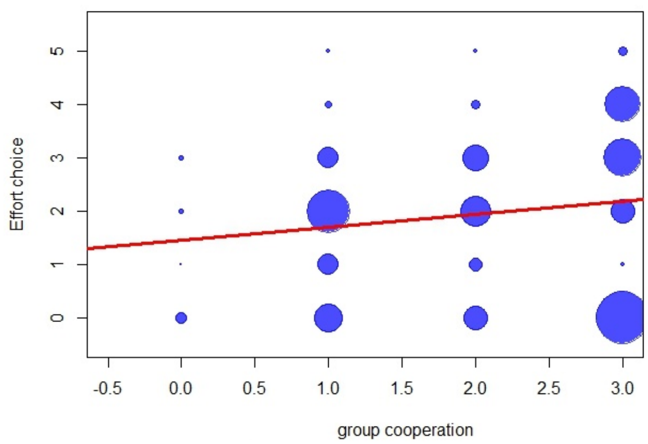
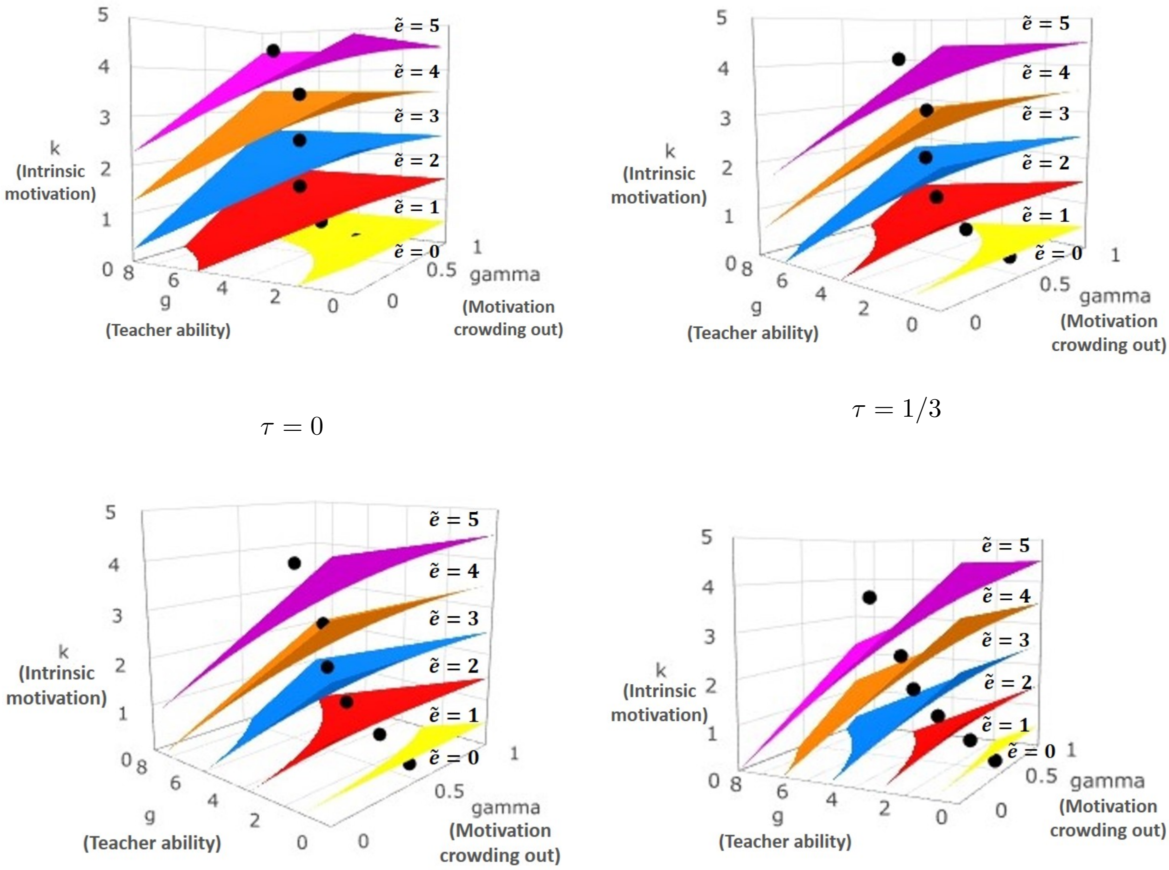


Figure 6: Joint frequency distribution of effort choice and reported group cooperation level in group incentive schools. The line represents the results from a linear regression of effort on reported group cooperation using all the observations. The correlation coefficient obtained from this is 0.24 with a p-value of 0.019.

For each of the points of support for the available observed measure of τ, the diagrams below represent the three-dimensional set for teacher heterogeneity consistent with each choice of effort. Plotting k on the z-axis and γ and g on the other axes, we can represent the parameter-space and the joint parameter sets that are consistent with each observed choice as illustrated below.



$\tau = 0$ $\tau = 1/3$

$\tau = 2/3$ $\tau = 1$

Figure 7: Sets for teacher heterogeneity in group incentive schools (Boundaries and Barycenters). The coloured surfaces indicate the separating boundaries of the sets. Barycenters of each set $E[\gamma, g, k|e]$, marked in black dots, were calculated based on uniformly drawing 10,000 values of $\{k, \gamma, g\}$ within this cuboid and averaging over the values of the parameters which fall in each set.

For each value of τ, the area below the yellow surface represents the parameter set for $\tilde{e} = 0$. The area between the red and yellow surfaces represents the parameter set for $\tilde{e} = 1$. The area between the blue and red surfaces represents the parameter set for $\tilde{e} = 2$. The area between the orange and blue surfaces represents the parameter set for $\tilde{e} = 3$. The area between the purple and orange surfaces represents the parameter set for $\tilde{e} = 4$. The area above the purple surface is consistent with choice $\tilde{e} = 5$.

As before, since the utility function is invariant to monotonic transformations, I set $\beta = 0.5$ in deriving the parameter sets below. The parameters α, which is the fixed salary, k, which represents intrinsic motivation, and the cost function are therefore to be interpreted as being scaled by a factor of 0.001.

Using simulations based on uniformly drawing 10,000 values of $\{k, \gamma, g\}$ within these cuboids, the barycenters $E[\gamma, g, k|e, \tau]$ are marked as black dots in the diagrams above. These can therefore be interpreted as indicative of the characteristics of the average teacher in each set. As in the individual incentives case, plotting the barycenters across the sets consistent with each effort choice shows an increasing trend in intrinsic motivation and ability as we move up the sets consistent with higher effort choices. The trend in motivation crowding out is decreasing for lower values of reported group cooperation τ but less clear and pronounced with higher levels of τ.[57]

5 Discussion

5.1 Counterfactual Policy Analysis and Policy Implications

Using the identified parameter sets for teacher heterogeneity in the preceding section in individual-level and group incentive schools, we can place bounds on the incentive intensity needed to increase a teacher's effort choice.

5.1.1 Individual-level incentives

The identified set for teacher heterogeneity consistent with an effort choice of zero in the preceding section was $\beta^{orig.}(1-\gamma_i)g_i + k_i \leqslant 0.5$ with $\beta^{orig.} = 0.5$. To increase a teacher's effort choice to one, we need $\beta^{new}(1-\gamma_i)g_i + k_i \geqslant 0.5$. By rearranging this expression as $\beta^{new} \geqslant \frac{0.5-k_i}{(1-\gamma_i)g_i}$ and substituting the minimum and maximum possible values of k_i from the identified set for effort of zero, one can see that a value of $\beta^{new} \in \left[0.5, \frac{0.5}{(1-\gamma_i)g_i}\right]$ would ensure that the teacher increases his/her effort choice from zero to one. The actual value of β^{new} needed in this set depends on the actual value of k_i of the teacher but the upper bound of this set, derived by taking the most pessimistic value of k_i consistent with the teacher's effort choice of zero under the original scheme, will be sufficient to push this teacher to increase his/her effort by at least one unit. Using similar reasoning, one can derive that a value of $\beta^{new} \in \left[0.5, 0.5 + \frac{1}{(1-\gamma_i)g_i}\right]$ would increase a teacher's effort choice by one if the teacher was choosing an effort between one and four inclusive under the original scheme with $\beta^{orig.} = 0.5$.

Thus to ensure all the teachers increased their effort choice by at least one unit when provided individual-level incentives, the bounds for the incentive intensity needed will be:

(a) for $\tilde{e}_i^{orig.} = 0 \rightarrow \tilde{e}_i^{new} \geqslant 1 \quad \forall i$

$$\beta^{new} \in \left[0.5, \max_{\{i:e_i^{orig.}=0|\beta^{orig.}=0.5\}} \left\{\frac{0.5}{(1-\gamma_i)g_i}\right\}\right]$$

[57] The section 7.5 in the appendix provides a detailed plot of the barycenters.

(b) for $\tilde{e}_i^{orig.} \neq 0 \rightarrow \tilde{e}_i^{new} \geqslant \tilde{e}_i^{orig.} + 1 \quad \forall i$

$$\beta^{new} \in \left[0.5, \max_{\{i:e_i^{orig.} \in [1,4] | \beta^{orig.}=0.5\}} \left\{0.5 + \frac{1}{(1-\gamma_i)g_i}\right\}\right]$$

One can make the following two observations:(1) the upper bound for the incentive intensity needed to increase a teacher's effort by one is decreasing in a teacher's ability and increasing in a teacher's motivation crowding out; and (2) increasing the effort of all teachers whose initial effort choice was non-zero is harder than that of teachers whose effort choice was zero when under the original incentive scheme with $\beta^{orig.} = 0.5$. The latter observation follows by noticing that the upper bound for β^{new} is higher in the case for increasing effort choice which was non-zero to begin with than that required for increasing effort choice from zero to one. Furthermore, given that the cost of increasing the teacher's effort choice by two or more units would be additive, one can note that the cost of increasing a teacher's effort from that chosen under the experiment is increasing in the increase in effort required.

Using the barycenters of the sets in Figure 4, we can go further and derive the incentive coefficient that would be needed to raise the effort choice by one of the average teacher choosing an effort level under the experimental incentive scheme. The table below provides the calculated incentive coefficients needed. We can observe that we need the incentive payment to be at least Rs. 730 to raise the effort of the average teacher choosing effort level of zero under the original incentive scheme by one unit. Furthermore, we need the incentive payment to be at least Rs. 1020 to increase the effort choice of the average teacher choosing a non-zero effort level under the original incentive scheme. Therefore, we need at least Rs. 1020 to increase the effort of all teachers not exerting the maximal effort level of 5 under the original incentive payment of Rs. 500 by one unit.

Table 2: Incentive intensity needed to raise the effort of the average teacher by one unit in individual-level incentive schools.

Effort	**Barycenters**			**β needed (Rs. 1000)**
	k	γ	g	
0	0.19	0.69	1.36	0.73
1	0.68	0.58	1.91	1.02
2	1.47	0.51	2.24	0.93
3	2.43	0.50	2.26	0.94
4	3.45	0.51	2.28	0.95
5	4.37	0.40	2.63	

Incentive coefficient needed to raise the effort choice of the average teacher by one in individual incentive schools. Barycenters of each set were calculated based on uniformly drawing 10,000 values of $\{k, \gamma, g\}$ within the cuboid containing the parameter sets in Figure 4 and averaging over the values of the parameters which fall in each set. k denotes intrinsic motivation, γ denotes the extent of motivation crowding out and g denotes teacher's ability. The final column calculates the values of β, the incentive intensity, needed to raise the average teacher's effort in each of the sets by one unit.

5.1.2 Group incentives

Using similar reasoning as in the individual-incentives case, one can derive that to increase all the teachers' effort choice by at least one unit under group incentives, the bounds for the incentive intensity needed are:

(a) for $\tilde{e}_i^{orig.} = 0 \rightarrow \tilde{e}_i^{new} \geqslant 1 \quad \forall i$

$$\beta^{new} \in \left[0.5, \max_{\{i: e_i^{orig.}=0 | \beta^{orig.}=0.5\}} \left\{ \frac{1}{(1-\gamma_i+\tau_i)g_i} \right\} \right]$$

(b) for $\tilde{e}_i^{orig.} \neq 0 \rightarrow \tilde{e}_i^{new} \geqslant \tilde{e}_i^{orig.} + 1 \quad \forall i$

$$\beta^{new} \in \left[0.5, \max_{\{i: e_i^{orig.} \in [1,4] | \beta^{orig.}=0.5\}} \left\{ 0.5 + \frac{2}{(1-\gamma_i+\tau_i)g_i} \right\} \right]$$

As in the individual-level incentives case, we can note that the upper bound for the incentive intensity needed to increase a teacher's effort by one is decreasing in a teacher's ability and increasing in a teacher's motivation crowding out. In addition, the upper bound is also decreasing in a teacher's reported group cooperation level. As before, by noticing that the upper bound in case (b) is stricter than that in case (a), increasing the effort of all teachers whose initial effort choice was non-zero is harder than that of teachers whose effort choice was zero when under the original incentive scheme with $\beta^{orig.} = 0.5$. As under individual-level incentives, given that the cost of increasing the teacher's effort choice by two or more units would simply cumulate accordingly, one can observe that the cost of increasing a teacher's effort from that chosen under the experiment is increasing in the increase in effort required.

Using the barycenters of the sets in Figure 7, the table below calculates the incentive coefficient needed to raise the effort choice of the average teacher under the experimental scheme by one.

We can observe that we need the incentive payment to be at least Rs. 720, Rs. 850, Rs. 900 and Rs. 1020 for respective values of τ of $\{0, 1/3, 2/3, 1\}$ to raise the effort of the average teacher choosing effort level of zero under the original incentive scheme by one unit.[58] Likewise, we need the incentive payment to be at least Rs. 990, Rs. 830, Rs. 810 and Rs. 860 for respective values of τ of $\{0, 1/3, 2/3, 1\}$ to increase the effort choice of the average teacher choosing a non-zero effort level under the original incentive scheme.

Therefore, depending on the reported group cooperation level of teachers, we need at least Rs. 720 - Rs. 1020 to increase the effort of all teachers not exerting the maximal effort level of 5 under the original incentive payment of Rs. 500 by one unit. Therefore to increase the effort levels of all the teachers not choosing the maximal level of effort under the original incentive scheme by at least one unit, we need the incentive payment to be

[58] The value of incentive payment for the average teacher needed is counter-intuitively increasing as τ increases because the values of $\{k, g\}$ at the barycenters for each set are decreasing as τ increases, completely offsetting and in fact, overturning the effect of increasing τ on β.

at least Rs. 850 - Rs. 1020 depending on the reported group cooperation levels of the teachers under the group incentive scheme.

Table 3: Incentive intensity needed to raise the effort of the average teacher by one unit in group incentive schools.

Effort	**Barycenters**			**β needed (Rs. 1000)**
	k	γ	g	
		$\tau = 0$		
0	0.19	0.69	2.81	0.72
1	0.67	0.55	3.75	0.99
2	1.45	0.51	4.48	0.96
3	2.45	0.50	4.46	0.94
4	3.45	0.51	4.52	0.95
5	4.37	0.42	5.23	
		$\tau = 1/3$		
0	0.16	0.59	1.09	0.85
1	0.56	0.58	3.05	0.83
2	1.21	0.53	4.17	0.78
3	2.08	0.49	4.43	0.77
4	3.07	0.51	4.51	0.77
5	4.15	0.45	5.42	
		$\tau = 2/3$		
0	0.16	0.55	0.68	0.90
1	0.54	0.56	2.15	0.81
2	1.05	0.55	3.66	0.71
3	1.78	0.52	4.34	0.69
4	2.71	0.50	4.40	0.69
5	3.91	0.46	5.56	
		$\tau = 1$		
0	0.17	0.53	0.44	1.02
1	0.55	0.53	1.50	0.86
2	1.00	0.56	2.97	0.70
3	1.59	0.54	3.98	0.66
4	2.36	0.50	4.38	0.65
5	3.68	0.47	5.62	

Incentive coefficient needed to raise the effort choice of the average teacher by one in group incentive schools. Barycenters of sets describing teacher heterogeneity consistent with each effort choice were calculated based on 10,000 simulations of values of $\{k, \gamma, g\}$ within the cuboid containing the parameter sets in Figure 7. k denotes intrinsic motivation, γ denotes the extent of motivation crowding out, g denotes teacher's ability and τ denotes the teacher's reported group cooperation level. The final column calculates the values of β, the incentive intensity, needed to raise the average teacher's effort in each of the sets by one unit.

6 Conclusion

This paper has outlined a simple theoretical framework which can be used to analyse performance related pay for teachers based on improvement in student test scores. Building on and going beyond the existing work in this field, I address the lacuna in the theoretical literature on the interaction of intrinsic and extrinsic motivation through explicit performance related pay. In light of the psychological and experimental literature on possible perverse and unintended consequences of explicit monetary rewards, I model teachers as motivated agents and allow for a motivation crowding out effect. The model highlights how the dimensions of heterogeneity among teachers - their ability, intrinsic motivation, extent to which they experience motivation crowding out due to explicit performance pay and internalisation of the externality of their effort on the other team members' earnings - and their interactions are important in determining teacher effort under incentives.

I then show how this framework can be used to explain the findings of the randomised control trial in India conducted by Muralidharan and Sundararaman (2011). In structural analysis, I refrain from making any untestable assumptions about the underlying distribution of unobservable heterogeneity or simply treating it as a multivariate error term. Instead, I bound the set of heterogeneous parameters using the necessary conditions that the set must satisfy if the choice behaviour were generated by the theoretical model, thereby adopting a set identification strategy.

Using the barycenters of the identified sets of teacher heterogeneity, I find that increasing the effort of the average teacher who did not change his/her effort under the individual incentive scheme, would need a higher bonus payment of Rs. 730. The same figure lies between Rs. 720 - Rs. 1020, depending on the reported group cooperation level among teachers, when providing group incentives. Furthermore, I find that increasing the change in effort of the teachers not choosing the maximal level of effort under the original payment scheme, would require an incentive payment of at least Rs. 1020 under individual-level incentives and at least Rs. 850 - Rs. 1020 under the group incentive payment scheme depending on the reported group cooperation level among the teachers.

I have argued that this simple static framework is sufficient to analyse the interaction between intrinsic and extrinsic motivation and to shed light on the tradeoff that the principal may face between incentive provision and motivation crowding out. It goes without saying, however, that extending the model to allow for dynamic contracting is an important next step and may help to further illuminate the findings of the experiment over time. Investigating how the optimal incentive intensity evolves as the teachers' reference points for effort change and how extrinsic and intrinsic motivation interact in the presence of dynamic incentives is an exciting research endeavour. The results derived from the model used here will nonetheless continue to hold for teachers who have a very low discount factor and/or assigned a low probability for the continuation of the program for the second year. Some interesting avenues and extensions of the analysis here would be to allow for psychic costs of effort as another dimension of heterogeneity, to compare teacher behaviour under non-monetary incentives and to vary the student test score technology to investigate the optimal effort level to incentivise under different scenarios.

References

[1] EFA Global Monitoring Report: Teaching and Learning. 2014.

[2] Dan Ariely, Anat Bracha, and Stephan Meier. Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially. *American Economic Review*, 99(1):544–555, 2009.

[3] Nava Ashraf, Oriana Bandiera, and Kelsey Jack. No margin no mission?: a field experiment on incentives for pro-social tasks. 2012.

[4] George Baker. Incentive Contracts and Performance Measurement. *The Journal of Political Economy*, 100(3):598, 1992.

[5] George Baker. The Use of Performance Measures in Incentive Contracting. *American Economic Review*, 90(2):415–420, 2000.

[6] Oriana Bandiera, Iwan Barankay, and Imran Rasul. Field Experiments with Firms. *Journal of Economic Perspectives*, 25(3):63–82, 2011.

[7] Oriana Bandiera, Iwan Barankay, and Imran Rasul. Team Incentives: Evidence from a Firm Level Experiment. *Journal Of The European Economic Association*, 11(5):pp1079–1114, 2013.

[8] Felipe Barrera-Osorio and Dhushyanth Raju. Teacher performance pay: Experimental evidence from Pakistan. *Journal of Public Economics*, 148:75–91, 2017.

[9] Jere R. Behrman, Susan W. Parker, Petra E. Todd, and Kenneth I. Wolpin. Aligning Learning Incentives of Students and Teachers: Results from a Social Experiment in Mexican High Schools. *Journal of Political Economy*, 123(2):325–364, 2015.

[10] Roland Benabou and Jean Tirole. Intrinsic and Extrinsic Motivation. *Review of Economic Studies*, 70(3):489–520, 2003.

[11] Roland Benabou and Jean Tirole. Incentives and prosocial behaviour. *IDEAS Working Paper Series from RePEc*, 2004.

[12] Roland Benabou and Jean Tirole. Bonus Culture: Competitive Pay, Screening, and Multitasking. *J. Polit. Econ.*, 124(2):305–370, 2016.

[13] Timothy Besley and Maitreesh Ghatak. Competition and Incentives with Motivated Agents. *American Economic Review*, 95(3):616–636, 2005.

[14] Timothy Besley and Maitreesh Ghatak. Sorting With Motivated Agents: Implications for School Competition and Teacher Incentives. *Journal of the European Economic Association*, 4(2-3):404–414, 2006.

[15] Timothy Besley and Maitreesh Ghatak. Status Incentives. *American Economic Review*, 98(2):206–211, 2008.

[16] Simon Burgess and Marisa Ratto. The Role of Incentives in the Public Sector Issues and Evidence. *Oxford Review of Economic Policy*, 19(2):285–300, 2003.

[17] Nazmul Chaudhury, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. Missing in Action: Teacher and Health Worker Absence in Developing Countries. *Journal of Economic Perspectives*, 20(1):91–116, 2006.

[18] Jacobus Cilliers, Ibrahim Kasirye, Clare Leaver, Pieter Serneels, and Andrew Zeitlink. Pay For Locally Monitored Performance - A Welfare Analysis for Teacher Attendance in Ugandan Primary Schools. *Economic Policy Research Centre*, 2016.

[19] Hai-Anh H. Dang and Elizabeth M. King. Incentives and teacher effort. *Economics of Transition*, 24(4):621–660, 2016.

[20] El Deci, R. Koestner, and R. M. Ryan. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol. Bull.*, 125:627–668, 1999.

[21] Thomas S. Dee and James Wyckoff. Incentives, Selection, and Teacher Performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, 34(2):267–297, 2015.

[22] Stefano DellaVigna and Devin Pope. What Motivates Effort? Evidence and Expert Forecasts. *Working Paper Series, National Bureau of Economic Research*, (22193), April 2016.

[23] Avinash Dixit. Incentives and organizations in the public sector: an interpretative review. *Journal of human resources*, 37(4):696–727, 2002.

[24] Esther Duflo and Rema Hanna. Monitoring Works: Getting Teachers to Come to School. 2005.

[25] Esther Duflo, Rema Hanna, and Stephen P. Ryan. Incentives Work: Getting Teachers to Come to School. *American Economic Review*, 102(4):1241–78, 2012.

[26] Tore Ellingsen and Magnus Johannesson. Paying Respect. *Journal of Economic Perspectives*, 21(4):135–149, 2007.

[27] Tore Ellingsen and Magnus Johannesson. Pride and Prejudice: The Human Side of Incentive Theory. *American Economic Review*, 98(3):990–1008, 2008.

[28] Armin Falk and Michael Kosfeld. The Hidden Costs of Control. *American Economic Review*, 96(5):1611–1630, 2006.

[29] Bruno S. Frey and Felix Oberholzer-Gee. The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out. *American Economic Review*, 87(4):746–755, 1997.

[30] Guido Friebel, Michael Kosfeld, and Gerd Thielmann. Trust the Police? Self-Selection of Motivated Agents into the German Police Force. 2016.

[31] A. Fuster and S Meier. Another Hidden Cost of Incentives: The Detrimental Effect on Norm Enforcement. *Management Science*, 56(1):57–70, 2010.

[32] Paul Glewwe, Nauman Ilias, and Michael Kremer. Teacher Incentives. *American Economic Journal: Applied Economics*, 2(3):205–227, 2010.

[33] Uri Gneezy, Stephan Meier, and Pedro Rey-Biel. When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, 25(4):191–210, 2011.

[34] Uri Gneezy and Aldo Rustichini. Pay Enough or Don't Pay at All. *The Quarterly Journal of Economics*, 115(3):791–810, 2000.

[35] Andrea Hammermann and Alwine Mohnen. The prize of hard work:Different incentive effects of non-monetary and monetary prizes: Different incentive effects of non-monetary and monetary prizes. *Journal of Economic Psychology*, 2014.

[36] Sarojini Hirshleifer. Incentives for Effort or Outputs? A Field Experiment to Improve Student Performance (Working Paper). 2015.

[37] B. Holmstrom. Moral Hazard in Teams. *Bell Journal of Economics*, 13(2):324, 1982.

[38] Bengt Holmstrom and Paul Milgrom. Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics & Organization*, 7:24–52, 1991.

[39] Hideshi Itoh. Cooperation in Hierarchical Organizations: An Incentive Perspective. *Journal of Law, Economics, Organization*, 8(2):321–345, 1992.

[40] C. K. Jackson. Match quality, worker productivity, and worker mobility: Direct evidence from teachers. *Review of Economics and Statistics*, 95(4):1096–1116, 2013.

[41] C. Kirabo Jackson and Elias Bruegmann. Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. *American Economic Journal: Applied Economics*, 1(4):85–108, 2009.

[42] Brian A Jacob and Lars Lefgren. Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education. *Journal of Labor Economics*, 26(1):101–136, 2008.

[43] Seema Jayachandran. Incentives to teach badly: After-school tutoring in developing countries. *Journal of Development Economics*, 108:190–205, 2014.

[44] Michael D. Jones. Teacher behavior under performance pay incentives. *Economics of Education Review*, 37:148–164, 2013.

[45] Naureen Karachiwalla and Albert Park. Promotion incentives in the public sector: Evidence from Chinese schools. *Journal of Public Economics*, 2016.

[46] M. P. Keane, P. E. Todd, and K. I. Wolpin. *The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications*, volume 4. 2011.

[47] Michael P. Keane. Structural vs. atheoretic approaches to econometrics. *Journal of Econometrics*, 156:3–20, 2010.

[48] Jaesoo Kim. Managerial beliefs and incentive policies. *Journal of Economic Behavior and Organization*, 119:84–95, 2015.

[49] Esteban F. Klor, Sebastian Kube, Eyal Winter, and Ro'i Zultan. Can higher rewards lead to less effort? Incentive reversal in teams. *Journal of Economic Behavior and Organization*, 97:72–83, 2014.

[50] Michael Kremer, Nazmul Chaudhury, F. Halsey Rogers, Karthik Muralidharan, and Jeffrey Hammer. Teacher Absence in India: A Snapshot. *Journal of the European Economic Association*, 3(2/3):658–667, 2005.

[51] Victor Lavy. Performance Pay and Teachers' Effort, Productivity, and Grading Ethics. *American Economic Review*, 99(5):1979–2011, 2009.

[52] Arthur Lewbel. The Identification Zoo – Meanings of Identification in Econometrics. *Journal of Economic Literature (forthcoming).*

[53] Alessandro Lizzeri, Margaret A Meyer, and Nicola Persico. The incentive effects of interim performance evaluations. 2002.

[54] Hamish Low and Costas Meghir. The Use of Structural Models in Econometrics. *Journal of Economic Perspectives*, 31(2):33–58, May 2017.

[55] Margaret A Meyer and John Vickers. Performance comparisons and dynamic incentives. *Journal of Political Economy*, 105(3):547–581, 1997.

[56] Karthik Muralidharan. Long-term Effects of Teacher Performance Pay: Experimental Evidence from India. *Society for Research on Educational Effectiveness*, 2012.

[57] Karthik Muralidharan and Venkatesh Sundararaman. Teacher Opinions on Performance Pay: Evidence from India. *Economics of Education Review*, 30(3):394–403, 2011.

[58] Karthik Muralidharan and Venkatesh Sundararaman. Teacher Performance Pay: Experimental Evidence from India. *Journal of Political Economy*, 119(1):39–77, 2011.

[59] Derek Neal. The Design of Performance Pay in Education. 2011.

[60] Susanne Neckermann, Reto Cueni, and Bruno S. Frey. Awards at work. *Labour Economics*, 31:205–217, 2014.

[61] Canice Prendergast. What Trade-off of Risk and Incentives? *American Economic Review*, 90(2):421–425, 2000.

[62] Canice Prendergast. Intrinsic Motivation and Incentives. *American Economic Review*, 98(2):201–205, 2008.

[63] Imran Rasul and Daniel Rogger. Management of bureaucrats and public service delivery : evidence from the Nigerian civil service. *Discussion paper (Centre for Economic Policy Research (Great Britain))*, 2016.

[64] Jonah E. Rockoff and Cecilia Speroni. Subjective and Objective Evaluations of Teacher Effectiveness. *American Economic Review*, 100(2):261–266, 2010.

[65] David E. M. Sappington. Incentives in Principal-Agent Relationships. *Journal of Economic Perspectives*, 5(2):45–66, 1991.

[66] Alexander Sebald and Markus Walzl. Optimal contracts based on subjective performance evaluations and reciprocity. *Journal of Economic Psychology*, 47:62–76, 2015.

[67] Douglas O. Staiger and Jonah E. Rockoff. Searching for Effective Teachers with Imperfect Information. *Journal of Economic Perspectives*, 24(3):97–118, 2010.

[68] Elie E. Tamer. Partial Identification in Econometrics. 2:167–195, 02 2010.

[69] Eric S. Taylor and John H. Tyler. The Effect of Evaluation on Teacher Performance. *American Economic Review*, 102(7):3628–3651, 2012.

[70] Ruth Wageman and George Baker. Incentives and cooperation: the joint effects of task and reward interdependence on group performance. *Journal of Organizational Behavior*, 18(2):139–158, 1997.

[71] Ludger Woessmann. Cross-country evidence on teacher performance pay. *Economics of Education Review*, 30(3):404–418, 2011.

[72] Andrew Zeitlin. Pay for Locally Monitored Performance? A Welfare Analysis for Teacher Attendance in Ugandan Primary Schools. 10118, 2016.

Data Source

The data used in the analysis here was taken from the data supplement to Karthik Muralidharan and Venkatesh Sundararaman, "Teacher Performance Pay: Experimental Evidence from India." Journal of Political Economy (2011): 39-77.
URL: http://www.journals.uchicago.edu/doi/suppl/10.1086/659655.

7 Appendix

7.1 Proofs

Proof of Corollary 1. Differentiating the first order condition obtained at the incentive compatible effort level for the intrinsically motivated with respect to β gives

$$\tilde{S}_1(e,\eta) + \beta\tilde{S}_{11}(e,\eta)\left(\frac{de}{d\beta}\right) + \phi_{11}(e,\beta)\left(\frac{de}{d\beta}\right) + \phi_{12}(e,\beta) = c''(e)\left(\frac{de}{d\beta}\right)$$

Rearranging gives: $\dfrac{de}{d\beta} = \dfrac{\tilde{S}_1(e,\eta) + \phi_{12}(e,\beta)}{c''(e) - \phi_{11}(e,\beta) - \beta\tilde{S}_{11}(e,\eta)}$

The result follows from noticing that the denominator is greater or equal to that of the self-interested agent for any given e as $\phi_{11}(e,\beta) \leqslant 0$ and $S_{11}(e,\eta) \leqslant 0$. The numerator is less than than that of a self-interested agent as $\phi_{12}(e,\beta) < 0$. □

Proof of Proposition 1. We have[59]

$$E[u_T|e] = \begin{cases} \alpha + \beta g\tilde{e} - \frac{1}{2}\tilde{e}^2 & for\ a\ self-interested\ agent \\ \alpha + \beta(1-\gamma)g\tilde{e} + ke - \frac{1}{2}\tilde{e}^2 & for\ a\ motivated\ agent \end{cases}$$

Using backward induction, the principal seeks to maximise its expected utility subject to meeting the participation and incentive compatibility constraints of the teacher. In the optimal contract, the value of α will be adjusted such that the participation constraint holds with equality.[60] Incentive compatibility will dictate:

$$\tilde{e} = \begin{cases} \beta g & for\ a\ self-interested\ agent \\ \beta(1-\gamma)g + k & for\ a\ motivated\ agent \end{cases}$$

[59]Since choosing e and choosing $\tilde{e}$ are equivalent, for notational convenience, I formulate a teacher's maximisation problem as one of choosing $\tilde{e}$.

[60]If a limited liability constraint is included in the analysis, which necessitates that the salary received by the teacher under this compensation structure should be at least as high as the subsistence level or a minimum wage, the optimal contract derived here is the one which holds when the limited liability constraint does not bind. This will be the case if the reservation utility of the teacher is high enough to result in an optimal base salary α that is greater than that which is needed to satisfy the limited liability constraint. I focus on this case here as a substantial wage premium currently prevails between government school teachers and the average per capita earnings in the Indian context; for instance, government school teachers' earnings are five times the average per-capita earnings in Andhra Pradesh (Muralidharan and Sundararaman, 2011). This may imply that their reservation utility is indeed quite high and in any case, taking this salary as their reference point, it is unlikely that the school teachers will be willing to accept contracts which force the base pay all the way down to any minimum wage or subsistence level such that the limited liability constraint binds. Nonetheless, for completeness, please refer to section 7.2 of the Appendix for the derivation of the optimal contract under the more general case with a limited liability constraint.

In line with Lemma 1, we have that if $k > \beta\gamma g$ so that $\phi_1(e, \beta) > 0$, then a motivated teacher's effort for a given β will be higher than that of a self-interested teacher. This case is illustrated in the figure 1 above where the optimal effort in green represents the effort level chosen by the motivated agent whilst that in red denotes the lower effort level chosen by a self-interested agent. In line with Lemma 2, we have that $\frac{d\tilde{e}}{d\beta} = g(1-\gamma)$ for the motivated agent and $\frac{d\tilde{e}}{d\beta} = g$ for the self-interested agent so that the responsiveness of the motivated agent to an increase in incentive intensity is less than that of the self-interested agent at the optimum.

The optimal intensity derived is:[61]

$$\beta^* = \begin{cases} \dfrac{f}{g} & for\ a\ self-interested\ agent \\ \max\left\{\dfrac{f(1-\gamma)-\gamma k}{g(1-\gamma^2)}, 0\right\} & for\ a\ motivated\ agent \end{cases}$$

One can now make the following observations:

(i) The optimal intensity for the motivated agent is decreasing in k in the presence of a motivation crowding out effect since

$$\frac{\partial\beta^*}{\partial k} = -\frac{\gamma}{g(1-\gamma^2)}$$

is negative for $\gamma \in (0, 1)$.

The optimal intensity for the motivated agent is also decreasing in γ since

$$\frac{\partial\beta^*}{\partial\gamma} = -\left(\frac{fg(1-\gamma)^2 + kg(1+\gamma^2)}{g^2(1-\gamma^2)^2}\right)$$

is negative for $\gamma \in [0, 1)$.

(ii) Further, the cross-partial derivative

$$\frac{\partial}{\partial\gamma}\left(\frac{\partial\beta^*}{\partial k}\right) = -\left(\frac{1+\gamma^2}{g(1-\gamma^2)^2}\right)$$

is negative for the values of γ under consideration.

(iii) The optimal incentive intensity parameter β^* for the motivated agent is bounded above by that of the self-interested agent. This can be seen more clearly if we rearrange terms in the optimal intensity for the motivated agent:

$$\beta^* = \frac{f(1-\gamma)-\gamma k}{g(1-\gamma^2)}$$

$$\Rightarrow\ \beta^* = \frac{f}{g(1+\gamma)} - \frac{\gamma k}{g(1-\gamma^2)}$$

[61] I assume that $\overline{u}$ is such that the principal obtains a non-negative payoff once the participation and incentive compatibility constraints are met. Hence, the model focuses on the determination of the optimal incentive intensity parameter. More precisely, $\overline{u} \leqslant \frac{(f(1-\gamma)+k)^2}{2(1-\gamma^2)+(f+k)e_0}$ and $\alpha \leqslant \frac{\gamma(f(1-\gamma)+k)^2}{(1-\gamma^2)(1+\gamma)} + fe_0$ is sufficient to ensure this.

One can now observe that the first term is lower than $\frac{f}{g}$ which is the optimal intensity parameter for a self-interested agent and the second term further decreases this value as the term is positive for the values of γ considered here. □

Proof of Corollary 2. We can observe that as $\gamma \to 0$, that is, when there is no motivation crowding out effect following the provision of monetary incentives, then $\beta^* \to \frac{f}{g}$ and $\tilde{e}^* \to \beta g + k \quad \Rightarrow \quad \tilde{e}^* \to f + k$ which is the first-best level of effort. However, if $\gamma \neq 0$, one can note that the effort achieved under the optimal contract derived above is $\tilde{e}^* = \frac{f(1-\gamma)+k}{1+\gamma}$ which is strictly less than the first-best $\tilde{e}^{**} = f + k$. □

Proof of Proposition 2. Using backward induction, the principal's problem is to maximise his expected payoff subject to the participation and incentive compatibility constraints for the two agents.[62]

$$\max_{\{\alpha_1,\alpha_2,\beta,\tilde{e}_1,\tilde{e}_2\}} E[u_G|e_1,e_2] = f(e_1+e_2) - \alpha_1 - \alpha_2 - \beta(g_1\tilde{e}_1 + g_2\tilde{e}_2)$$

s.t.

$$E[u_{T_i}|\mathbf{e}] \geqslant \overline{u} \quad i = 1,2 \qquad (PC_i)$$

$$\tilde{e}_i = \frac{\beta(1-\gamma_i+\tau_i)g_i}{2} + k_i \quad i = 1,2 \quad (IC_i)$$

Substituting the two participation constraints in the principal's objective function, reduces the problem to[63]

$$\max_{\{\beta,\tilde{e}_1,\tilde{e}_2\}} E[u_G|\mathbf{e}] = f(e_1+e_2) - \frac{\beta}{2}(\gamma_1 g_1\tilde{e}_1 + \gamma_2 g_2\tilde{e}_2) + k_1e_1 + k_2e_2 - \frac{1}{2}(\tilde{e}_1^2 + \tilde{e}_2^2) - 2\overline{u}$$

s.t.

$$\tilde{e}_i = \beta\frac{(1-\gamma_i+\tau_i)g_i}{2} + k_i \quad i = 1,2 \quad (IC_i)$$

Solving this yields[64]

$$\beta^* = \max\left\{ \frac{f((1+\tau_1-\gamma_1)g_1 + (1+\tau_2-\gamma_2)g_2) - (\gamma_1k_1g_1 + \gamma_2k_2g_2)}{\frac{g_1^2}{2}((1+\tau_1)^2-\gamma_1^2) + \frac{g_2^2}{2}((1+\tau_2)^2-\gamma_2^2)}, 0 \right\}$$

(i) One can note that the optimal incentive intensity is decreasing in a teacher's intrinsic motivation k_i as the numerator is always decreasing in k_i if the teacher experiences motivation crowding out whilst the denominator is positive and independent of a teacher's intrinsic motivation.

[62] As before, since $\tilde{e} = e - e_0$ where e_0 is predetermined at the time of choosing an effort level e, choosing e and choosing $\tilde{e}$ are equivalent. I therefore formulate the principal's problem as one of choosing $\tilde{e}$ for both teachers.

[63] α in each case is chosen to satisfy the participation constraint of the agent.

[64] Given the concavity of the objective function, the first order condition is sufficient for a maximum.

(ii) The optimal incentive intensity for the case when $\{\overline{\gamma}, \overline{\tau}\} = (0,0)$ reduces to $\beta^*_{\overline{\gamma}=0,\overline{\tau}=0} = \dfrac{2f(g_1+g_2)}{g_1^2+g_2^2}$.

For $\{\overline{\gamma}, \overline{\tau}\} = (0,1)$, it is $\beta^*_{\overline{\gamma}=0,\overline{\tau}=1} = \dfrac{f(g_1+g_2)}{g_1^2+g_2^2}$.

If $\{\overline{\gamma}, \overline{\tau}\} = (1,1)$, it is $\beta^*_{\overline{\gamma}=1,\overline{\tau}=1} = \max\left\{\dfrac{2}{3}\dfrac{f(g_1+g_2)-(g_1k_1+g_2k_2)}{g_1^2+g_2^2}, 0\right\}$.

If $\{\overline{\gamma}, \overline{\tau}\} = (1,0)$, then $\beta^*_{\overline{\gamma}=1,\overline{\tau}=0} = 0$.

One can observe that $\beta^*_{\overline{\gamma}=0,\overline{\tau}=0} = 2\beta^*_{\overline{\gamma}=0,\overline{\tau}=1}$ and $\beta^*_{\overline{\gamma}=1,\overline{\tau}=1} \leqslant \dfrac{2}{3}\beta^*_{\overline{\gamma}=0,\overline{\tau}=1}$. Hence the ordering of optimal incentive intensity in these different scenarios is as stated in the proposition.

Substituting the optimal incentive intensity in the incentive compatibility constraint for each agent derived in lemma 3, we can obtain the following results.

$$\tilde{e}^*_{i,\{\overline{\gamma}=0,\overline{\tau}=0\}} = \beta^*_{\overline{\gamma}=0,\overline{\tau}=0}\frac{g_i}{2} + k_i = \beta^*_{\overline{\gamma}=0,\overline{\tau}=1} g_i + k_i$$

where the last equality follows from noticing that $\beta^*_{\overline{\gamma}=0,\overline{\tau}=0} = 2\beta^*_{\overline{\gamma}=0,\overline{\tau}=1}$.

$$\tilde{e}^*_{i,\{\overline{\gamma}=0,\overline{\tau}=1\}} = \beta^*_{\overline{\gamma}=0,\overline{\tau}=1} g_i + k_i$$

$$\tilde{e}^*_{i,\{\overline{\gamma}=1,\overline{\tau}=1\}} = \beta^*_{\overline{\gamma}=1,\overline{\tau}=1}\frac{g_i}{2} + k_i$$

and given that $\beta^*_{\overline{\gamma}=1,\overline{\tau}=1} \leqslant \dfrac{2}{3}\beta^*_{\overline{\gamma}=0,\overline{\tau}=1}$, we can see that

$$\tilde{e}^*_{i,\{\overline{\gamma}=1,\overline{\tau}=1\}} \leqslant \frac{1}{3}\beta^*_{\overline{\gamma}=0,\overline{\tau}=1} g_i + k_i$$

And $\tilde{e}^*_{i,\{\overline{\gamma}=1,\overline{\tau}=0\}} = k_i$ since $\beta^*_{\overline{\gamma}=1,\overline{\tau}=0} = 0$.

The ordering follows immediately from comparing these effort levels.

(iii) By differentiating the optimal intensity with respect to γ_i, one can note that $g_{-i}(1+\tau_{-i}+\gamma_{-i}) \geqslant 2\gamma_i g_i$ is a sufficient condition to ensure that the optimal incentive intensity is decreasing in the teacher's intrinsic motivation. If the teachers are similar in their abilities and the extent to which they experience motivation crowding out, then this condition reduces to $1+\tau_i \geqslant \gamma$ which is always true as $\gamma \leqslant 1$. □

Proof of Corollary 5. The effort level when $\{\overline{\gamma}, \overline{\tau}\} = (0,1)$ is $\tilde{e}^*_{i,\{\overline{\gamma}=0,\overline{\tau}=1\}} = \beta^*_{\overline{\gamma}=0,\overline{\tau}=1} g_i + k_i$. Substituting for the optimal incentive intensity gives

$$\tilde{e}^*_{i,\{\overline{\gamma}=0,\overline{\tau}=1\}} = f\left(\frac{g_i^2+g_ig_{-i}}{g_i^2+g_{-i}^2}\right) + k_i$$

The first best effort level for each teacher is $\tilde{e}^{**}_i = f + k_i$ as derived in the benchmark case. One can observe that if the two teachers are identical in terms of their ability, then the incentive intensity reduces to $\dfrac{f}{g}$ and $\tilde{e}^*_i = \tilde{e}^{**}_i$. If however the two teachers differ in terms of their ability then, for the less able teacher, that is, for whom $g_i < g_{-i}$, we have that $f\left(\dfrac{g_i^2+g_ig_{-i}}{g_i^2+g_{-i}^2}\right) < f$.

Hence, the effort of the less able teacher in the team is lower than the first best effort level case. The opposite holds true for the more able teacher in the team. The sum of their efforts is

$$\tilde{e}^*_{i,\{\overline{\gamma}=0,\overline{\tau}=1\}} + \tilde{e}^*_{-i,\{\overline{\gamma}=0,\overline{\tau}=1\}} = f\left(\frac{(g_i+g_{-i})^2}{g_i^2+g_{-i}^2}\right) + k_i + k_{-i}$$

Since $\frac{(g_i+g_{-i})^2}{g_i^2+g_{-i}^2} < 2$ if $g_i \neq g_{-i}$, the sum of efforts achieved is strictly less than that under the first best case. □

7.2 Provision of incentives with a limited liability constraint

This section derives the optimal intensity for the more general case where a limited liability constraint may apply when effort being unobservable and non-contractible. Reasons for this may include minimum wages, a minimum base pay set by the government for all government employees or the requirement to give a minimum subsistence pay. I shall denote this here as $\underline{\alpha}$. In deriving the optimal contract under a limited liability constraint, I closely follow the steps set out in the appendix of Besley and Ghatak (2005).

The principal's problem is to maximise the expected payoff subject to a participation, limited liability and and incentive compatibility constraint denoted PC, LL and IC respectively.

$$\max_{\alpha,\beta,e} fe - \alpha - \beta g\tilde{e}$$

s.t.

$$\alpha + \beta(1-\gamma)g\tilde{e} + ke - \frac{1}{2}\tilde{e}^2 \geqslant \overline{u} \; (PC)$$

$$\alpha \geqslant \underline{\alpha} \qquad (LL)$$

$$\tilde{e} = \beta(1-\gamma)g + k \qquad (IC)$$

Substituting IC in the principal's objective function and PC, we get[65]:

$$\max_{\alpha,\beta}(f-\beta g)(\beta(1-\gamma)g+k) - \alpha + (f)e_0$$

s.t.

$$\alpha + ke_0 + \frac{1}{2}(\beta(1-\gamma)g+k)^2 \geqslant \overline{u} \; (PC)$$

$$\alpha \geqslant \underline{\alpha} \qquad (LL)$$

Lemma 1. *At least one of the participation or limited liability constraints will bind.*

[65] Since I replace e with $\tilde{e}$ in the principal's objective function and the agent's utility function when substituting in the IC, there will be an additional fe_0 and ke_0 term in the principal's and agent's payoff respectively.

Proof. Suppose instead that both PC and LL do not bind. Then the optimal intensity derived from the unconstrained maximisation is

$$\beta^* = \max\left\{\frac{f(1-\gamma)-k}{2(1-\gamma)g}, 0\right\}$$

Since by assumption LL and PC are not binding, the principal can optimally choose to lower α without affecting the agent's incentives and raise his payoff. $\square$

Lemma 2. *If PC binds but LL does not, we achieve the second-best case effort level as derived in the main body of the paper.*[66]

Lemma 3. *If PC does not bind but LL does i.e $\alpha = \underline{\alpha}$, then there are two possible cases for optimal intensity depending on whether $f(1-\gamma)$ is greater than k. Again, only the second-best effort level is achievable.*

Proof. As derived in Lemma 1 here, optimal intensity with PC not binding and $\alpha = \underline{\alpha}$ will be given by:

$$\beta^* = \max\left\{\frac{f(1-\gamma)-k}{2(1-\gamma)g}, 0\right\}$$

Case 1: $f(1-\gamma) \leqslant k$

$$\beta^* = 0$$
$$\tilde{e}^* = k$$
$$u_T = \underline{\alpha} + ke_0 + \frac{1}{2}k^2$$

Since by assumption PC does not bind, we need $\frac{1}{2}k^2 + ke_0 > \overline{u} - \underline{\alpha}$.

$$u_G = fk - \underline{\alpha} + fe_0$$

Hence we need $\underline{\alpha} \leqslant fk + fe_0$ for the principal to obtain a non-negative payoff here.

Case 2: $f(1-\gamma) > k$

$$\beta^* = \frac{f(1-\gamma)-k}{2(1-\gamma)g}$$
$$\tilde{e}^* = \frac{f(1-\gamma)+k}{2}$$
$$u_T = \underline{\alpha} + ke_0 + \frac{1}{8}(f(1-\gamma)+k)^2$$

Since by assumption PC does not bind, we need $ke_0 + \frac{1}{8}(f(1-\gamma)+k)^2 > \overline{u} - \underline{\alpha}$.

$$u_G = \frac{(f(1-\gamma)+k)^2}{4(1-\gamma)} - \underline{\alpha} + fe_0$$

Hence we need $\underline{\alpha} \leqslant \frac{(f(1-\gamma)+k)^2}{4(1-\gamma)} + fe_0$ for the principal to obtain a non-negative payoff here.

Note that in both cases $\tilde{e}^* < f + k$ which is the first-best level of effort. $\square$

[66] The principal obtains a non-negative payoff provided $\overline{u} \leqslant \frac{(f(1-\gamma)+k)^2}{2(1-\gamma^2)} + (f+k)e_0$ and $\underline{\alpha} < \alpha^* \leqslant \frac{\gamma(f(1-\gamma)+k)^2}{(1-\gamma^2)(1+\gamma)} + fe_0$.

Lemma 4. *When both PC and LL bind, the constraints uniquely determine the optimal α and β^*.*

Proof. By assumption, we have that $\alpha = \underline{\alpha}$ and $\underline{\alpha} + ke_0 + \frac{1}{2}(\beta(1-\gamma)g + k)^2 = \overline{u}$. Rearranging PC, we get that the optimal intensity

$$\beta^* = \frac{\sqrt{2(\overline{u} - \underline{\alpha} - ke_0)} - k}{(1-\gamma)g}$$

$$\tilde{e}^* = \sqrt{2(\overline{u} - \underline{\alpha} - ke_0)}$$

$$u_T = \overline{u}$$

$$u_G = \left[f - \frac{\sqrt{2(\overline{u} - \underline{\alpha} - ke_0)} - k}{1-\gamma} \right] (\sqrt{2(\overline{u} - \underline{\alpha} - ke_0)}) - \underline{\alpha} + fe_0$$

Hence we need $\underline{\alpha} \leqslant \left[f - \frac{\sqrt{2(\overline{u} - \underline{\alpha} - ke_0)} - k}{1-\gamma} \right] (\sqrt{2(\overline{u} - \underline{\alpha} - ke_0)})$ for the principal to obtain a non-negative payoff. □

Proposition 1. *The effort level achieved under the optimal contract when effort is unobservable and non-contractible is always lower than the first-best case in the presence of motivation crowding out. Optimal intensity is characterised by:*

$$\beta^* = \begin{cases} \max\left\{ \dfrac{f(1-\gamma) - k}{2(1-\gamma)g}, 0 \right\} & if\ \ \overline{u} \in [0, \underline{v}] \ \ and\ \ \alpha = \underline{\alpha} \\ \dfrac{\sqrt{2(\overline{u} - \underline{\alpha} - ke_0)} - k}{(1-\gamma)g} & if\ \ \overline{u} \in [\underline{v}, \overline{v}] \ \ and\ \ \alpha = \underline{\alpha} \\ \max\left\{ \dfrac{f(1-\gamma) - \gamma k}{g(1-\gamma^2)}, 0 \right\} & if\ \ \overline{u} \in [\underline{v}, \overline{v}] \ \ and\ \ \alpha > \underline{\alpha} \end{cases}$$

where $\underline{v}$ denotes all such outside options for which PC does not bind and $\underline{v}$ denotes the value of the outside option which drives principal's expected payoff to zero.

7.3 Test score distribution in treatment and control group schools

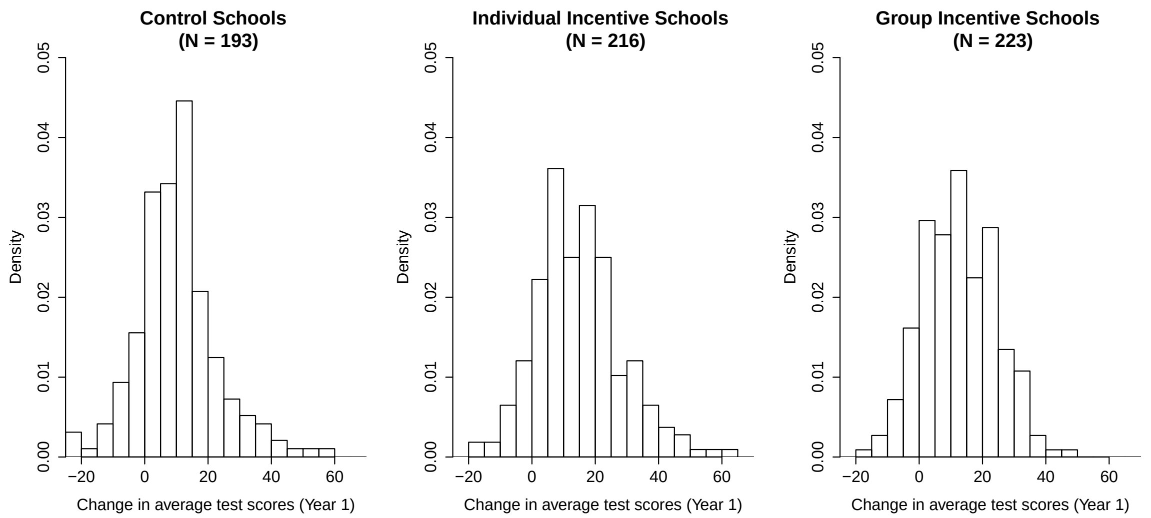


Figure 8: Changes in student test scores are reported in percentage point terms. N refers to the number of teachers.

7.4 Group cooperation in group incentive schools

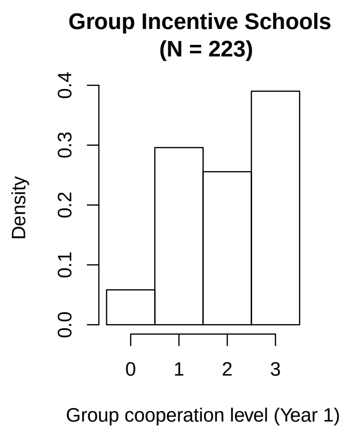


Figure 9: Histogram of group cooperation level in group incentive schools. The index aggregates the binary responses to questions asking (1) if the teachers shared the work load; (2) if they encouraged hard work; (3) if they shared teaching methods.

7.5 Barycenters for sets of teacher heterogeneity with group incentives

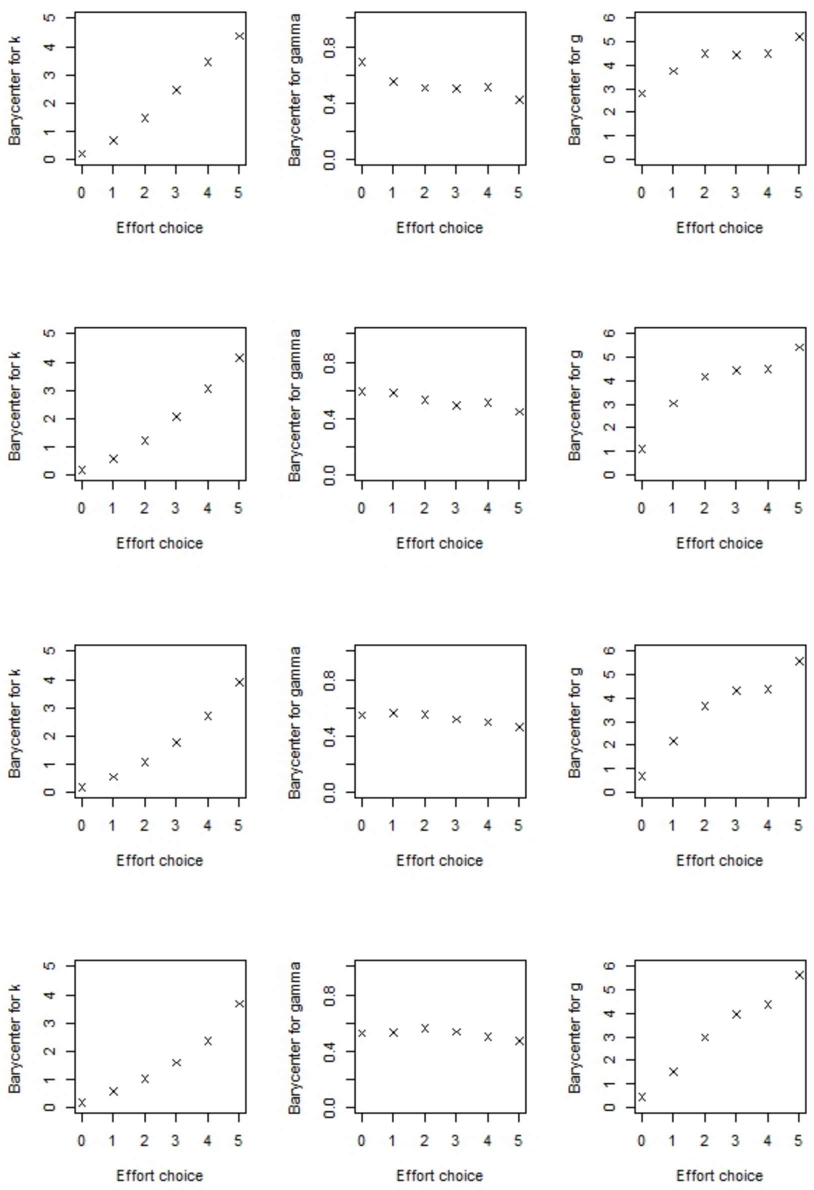


Figure 10: Barycenters for sets of teacher heterogeneity consistent with each effort choice in group incentive schools. Calculations based on 10,000 simulations of $\{k, \gamma, g\}$ within the cuboids of teacher heterogeneity and averaging over the values of the parameters which fall in each set.