Pay for locally monitored performance?

A welfare analysis for teacher attendance in Ugandan primary schools

Jacobus Cilliers,[*] Ibrahim Kasirye,[†] Clare Leaver,[‡]
Pieter Serneels,[§] and Andrew Zeitlin[¶]

January 2016

Abstract

To improve service delivery public sector organizations often rely on reports from interested parties that are costly to verify. If accurate, this information can then serve the dual purpose of incentivizing performance and reducing policy mistakes. Received wisdom suggests that the benefit of raising performance via financial incentives should be balanced against the risk of collusion between agent and monitor, resulting in false reporting. We evaluate two different forms of local monitoring by head-teachers of Ugandan primary schools, randomly varying whether reports of teacher attendance trigger financial incentive payments or not. A theoretical model provides a normative framework to make welfare comparisons, taking into account teacher attendance, the cost of monitoring, as well as policy mistakes due to false reporting. Consistent with the model, we find that teacher attendance, monitoring frequency, and the number of false reports all increased with the introduction of financial incentives. More surprisingly, but again consistent with the theory, we find that the number of policy mistakes actually decreases with incentives: there were more false reports but this effect was counter-balanced by more reports in general and (hence) fewer mistakes caused by a lack of information. We combine empirical estimates with the theoretical framework to bound welfare gains from attaching financial incentives to local reports, and conclude that welfare was higher with financial incentives.

[*]Georgetown University
[†]Economic Policy Research Centre
[‡]BSG University of Oxford
[§]University of East Anglia
[¶]Georgetown University

1 Introduction

To build incentives for civil servants, public sector organizations often rely on reports from interested parties that are costly to verify. These systems fail when managers or the agents themselves exert effort to game—and even to outright misreport—these performance metrics. When they go awry, such incentive schemes risk failure on two margins: they may fail to incentivize the desired behavior, and they may provide systematically biased information about outcomes that can mislead broader planning processes.

Public policy is replete with cases of delegated monitoring and incentive schemes gone wrong. Secretary of the U.S. Department of Veteran's affairs resigned after an audit revealed evidence of fraudulent behavior in 64 percent of VA health care facilities, where employees kept veterans off of official waiting lists in order to meet targeted 14-day waiting times for appointments.[1] Teachers have been found to cheat on high-stakes student assessment in the Atlanta, Chicago, and Washington, DC school districts.[2] French colonial history provides a dramatic tale of incentives gone wrong, when colonial administrators' attempts to incentivize the collection of rats resulted in widespread rat farming (Vann 2003). More recently, conditional transfers of cash or goods in kind have increasingly been used to increase demand for public services, but recent work has raised questions about agents' scope for manipulation of the metrics by which these conditions are measured (Linden & Shastry 2012).

In this paper, we address these issues in the context of Ugandan primary education. We ask whether and how incentives for civil servants—here, primary school teachers—can be effectively designed when based on *delegated* monitoring. This is an important issue in Ugandan education, where teacher absenteeism levels are such that pupils in rural, northern Uganda receive only 50 effective days of instruction in the entire school year (Wane & Martin 2013).[3] Beyond the direct advantage of incentivising attendance, the reported information can also be used by government for planning purposes. Reliable reporting is valuable over and above its instrumental use in inducing a behavioral response.

We develop an intuitive theoretical framework to understand how the preferences of both agents and monitors affects reporting and attendance, and how this depends on the financial stakes attached to the reports. This allows us to evaluate the joint welfare consideration on both metrics: attendance and quality of information. We then conduct a randomized, controlled trial that tests alternative monitoring schemes, experimentally varying the existence of financial stakes attached to local reports. This experiment provides tests of the predictions of the theory, and our theory provides a normative lens through which to gauge the welfare impacts of alternative intervention designs.

The theoretical model considers the the welfare of a teacher who decides whether to attend or not; a head-teacher who decides whether to monitor and what to report; a bureaucrat who makes a policy decision based on the report; and parents who care about teacher presence. Prior to these decisions the teacher and head-teacher bargain over their respective strategies. Under a regime of financial incentives

[1] See Bronstein and Griffin, "A fatal wait: Veterans languish and die on a VA hospital's secret list", CNN, April 23, 2014, online: http://www.cnn.com/2014/04/23/health/veterans-dying-health-care-delays/, and Veterans Health Administration (2014).

[2] On the Atlanta scandal, see Motoko Rich, "Scandal in Atlanta Reignites Debate Over Tests' Role", *New York Times*, April 2, 2013, online: http://www.nytimes.com/2013/04/03/education/atlanta-cheating-scandal-reignites-testing-debate.html. Problems in Chicago are discussed by Jacob and Levitt (2003).

[3] Comparable problems exist in schooling systems across the developing world (Chaudhury et al. 2006)

the teacher can receive a bonus, but only if the head-teacher makes a (costly) report and reports the teacher as present. The addition of a bonus grants the head-teacher, who is now also more willing to monitor, bargaining power to exert higher teacher attendance. Both monitoring and attendance thus increase with financial incentives. In turn, the bureaucrat makes a policy mistake whenever the head-teacher falsely reports an absent teacher as present. Crucially, the model also demonstrates that the bureaucrat makes a policy mistake when there is no report of an absent teacher. The theory thus predicts that policy mistakes due to false reports of presence will be higher with financial incentives, but policy mistakes due to non-reports will be lower, because both teacher presence and the frequency of monitoring is lower. The net welfare impact, however, is ambiguous.

Taking the theory to the data, we find that most predictions of the model hold: With the introduction of bonus payments, teacher attendance is higher and policy mistakes due to non-reports of absent teachers are lower. The broad support of the model provides confidence to use the theoretical framework to make welfare comparisons by empirically estimating the parameter values of the model. Given reasonable assumptions of parents preferences and cost of policy mistakes, we find that the government bureaucracy's expected welfare is higher in the case of financial incentives: Both teacher presence increased and policy mistakes were reduced.

This paper builds on a related stream of literature that considers monitoring and incentive schemes to improve public-sector outcomes, in developing-country education systems and beyond. In education, Muralidharan and Sundararaman (Muralidharan & Sundararaman 2011) demonstrate the potential effects of using measures of learning outcomes to incentivize teachers; however, such schemes may be costly to administer and may lead to distortionary activities by teachers (Glewwe, Ilias & Kremer 2010). Incentivizing teacher *inputs* is attractive insofar as these (a) margins of effort such as presence are both important and well aligned with the production of learning, and (b) these expose teachers to substantially less risk than test-based accountability. While 'automated' measurement of teacher presence has been effective in NGO schools in India (Duflo, Hanna & Ryan 2012), efforts to scale up such approaches among public-sector (health) workers were met with challenges (Banerjee et al. 2007). This suggests that understanding how monitoring schemes interact with the preferences of agents and monitors is crucial to designing scalable incentive schemes. This message is reinforced in recent work by Duflo and colleagues (forthcoming), who show that environmental auditors suffer from conflicts of interest that substantially affect the data they report and the subsequent behavioral responses of firms.

This paper makes three principal contributions to this literature. First, the randomized, controlled trial we conduct provides direct evidence of the efficacy of alternative policies to promote teacher presence in Uganda. Second, the theory of delegated monitoring developed and tested here has application to a variety of contexts—it may apply wherever incentives for public servants depend on reports by interested parties, with costly verification. The existing literature has highlighted a variety of challenges arising when policymakers pay for measures of performance, ranging from multitask problems (Holmstrom & Milgrom 1991) to the potential crowding out of intrinsic motivation (Deci 1971; Bénabou & Tirole 2003, 2006) to the potential for gaming (verifiable) performance metrics. Our theoretical framework highlights the welfare cost attributable to the distortion of administrative data that arises when delegated monitoring goes wrong. It also reveals the potential for an *improvement* in the quality information, despite risk of collusion, due to more frequent monitoring and higher overall performance (in our case, teacher attendance).

Third, from a methodological perspective, our paper demonstrates several advantages of closely incorporating theory in such experimental work: theory yields several testable predictions that we take to our data; it provides a lens to guide the selection and specification of outcomes considered; and it provides a normative framework through which welfare comparisons can be made across alternative treatment regimes.

The remainder of the paper proceeds as follows. Section 2 introduces a theory of delegated monitoring, which yields both positive predictions for monitors' and teachers' responses to alternative treatment regimes and normative, welfare criteria for comparing outcomes under each. Section 3 outlines the experiment and data. Section 4 takes the positive predictions of the model to the data. Section 5 considers the implications for welfare, using the theoretical framework to show how estimated enrollment impacts can be combined with alternative assumptions about the persistence of the intervention's impacts to guide policymakers' decisions between alternative monitoring regimes. Section 6 concludes.

2 Theory

The model provides a stylised way to analyse the effect of local monitoring and incentives on teacher attendance. The basic structure is that teachers can choose between showing up for work or not. In the standard context teacher attendance remains unmonitored, yielding a certain probability of teacher attendance. We then study how introducing local monitoring by the head teacher, who reports teacher attendance to a bureaucrat, affects teacher attendance. Next, we investigate how combining local monitoring and incentives triggered by the monitoring report, affect teacher attendance, and compare expected attendance across the three contracts. We then analyse monitoring failures and false reports to assess the probability of policy mistakes by bureaucrats. This leads to testable predictions that we later take to the data. Finally, the model provides a structure to carry out welfare analysis.

2.1 Model

Players and actions The economy consists of: a teacher (he), a head-teacher (she), a government bureaucracy (it), and n identical parents. In all arms, the teacher chooses whether to attend school, $a \in \{0, 1\}$. In the control arm, the head-teacher plays no active role. In both treatment arms, the head-teacher chooses whether to monitor $m \in \{0, 1\}$. If the head-teacher monitors, $m = 1$, she observes a and chooses a public report $r \in \{0, 1\}$. We will say that the head-teacher send a *truthful report* iff $r = a$. If the head-teacher does not monitor, $m = 0$, she cannot send a report.[4] In all arms, the bureaucracy takes a policy decision $p \in \{0, 1\}$. We will say that the bureaucracy makes a *policy mistake* whenever $p \neq a$. In the Info&Bonus treatment arm, the bureaucracy pays a cash bonus β directly to the teacher iff he is reported present, $r = 1$. Parents play no active role under any treatment arm.

[4]We provide evidence to support this assumption that monitoring is necessary to send a report in Section X below.

Payoffs All players are risk neutral. Net of any side-transfers, the payoffs of the strategic players are:

$$\begin{aligned} U^T &= 1_{\{m=1,r=1\}} \cdot \beta - 1_{\{a=1\}} \cdot C^T - 1_{\{r=0\}} \cdot \delta \\ U^H &= 1_{\{a=1\}} \cdot \varepsilon^H - 1_{\{m=1\}} \cdot C^H \\ U^G &= 1_{\{a=1\}} \cdot n\, \varepsilon^P - 1_{\{m=1,r=1\}} \cdot \beta - 1_{\{p\neq a\}} \cdot \kappa. \end{aligned}$$

If the teacher attends school, $a = 1$, he incurs a (possibly negative) cost of C^T. If the teacher attends, the head-teacher and a representative parent receive private benefits of ε^H and ε^P. If the head-teacher monitors, she incurs a (possibly negative) cost of C^H. Reporting entails no further cost for the head-teacher but a mark of absent, $r = 0$, imposes a reputational cost of δ on the teacher. If the head-teacher is indifferent, we assume that she reports truthfully. Finally, a policy mistake entails a loss of κ for the bureaucracy. We will use the bureaucracy's payoff−i.e. the total parental benefit from teacher attendance less the cost of cash bonus and/or policy mistake−as our welfare criterion.

Key Assumptions The costs C^T and C^H are observed by the head-teacher and teacher but not by any other player. From the bureaucracy's perspective, these costs are random variables. For convenience, we assume that realizations are drawn independently from uniform distributions. The lower and upper support of the distribution of C^T are denoted by $\underline{C}^T$ and $\overline{C}^T$. To calibrate the model to the baseline absenteeism rate, we assume $-\underline{C}^T > \overline{C}^T > 0$. The lower and upper support of the distribution of C^H are denoted similarly, although here we simply assume $\overline{C}^H > 0 > \underline{C}^H$. Again for convenience and in the spirit of rationalizing baseline absenteeism, we assume that the bonus β is the only source of transferable utility.[5] Relatedly, we assume that side-contracts sharing β are costless and enforceable, and that the head-teacher can commit to monitor. Finally, we assume that parameters satisfy

$$n\, \varepsilon^P > \frac{\varepsilon^H(\overline{C}^T - \underline{C}^T)}{\varepsilon^H - \delta} > \varepsilon^H > \delta > 0, \tag{1}$$

implying that the parental gain from teacher attendance must be sufficiently high.[6]

Timing To emphasize the differences across arms, it is worth spelling out the order of play. The timing in the Control arm is:

0. Nature draws a realization of C^T and reveals this cost to the teacher.

1. The teacher chooses whether to attend school, $a \in \{0, 1\}$. Simultaneously, the bureaucracy chooses a policy decision $p \in \{0, 1\}$. Payoffs are realized and the game ends.

The timing in the Info arm is:

0. The bureaucracy announces the monitoring scheme. Nature draws realizations of C^T and C^H and reveals both of these costs to the teacher and the head-teacher.

[5] If all sources of utility were transferable, then (via Coasian logic) the players should reach a jointly-efficient outcome. The high rates of absenteeism reported above suggest this is not the case. In reality, δ might be partly transferable. We assume non-transferability to make the distinction between the Info and Info&Bonus arms as clear as possible.

[6] As we discuss below, this is a sufficient condition for the bureaucracy to choose $\beta \geq \varepsilon^H$ in the Info&Bonus treatment.

1. The head-teacher chooses whether to make an announcement to the teacher. An announcement $R(a)$ commits the head-teacher to monitor, $m = 1$, and specifies the report r that the head-teacher will send to all players following each possible action a.

2. The teacher chooses whether to attend school, $a \in \{0, 1\}$.

3. If the head-teacher made the announcement at Stage 1, she monitors and sends the public report $r = R(a)$. Otherwise, the head-teacher takes no action.

4. The bureaucracy chooses a policy decision $p \in \{0, 1\}$. Payoffs are realized and the game ends.

The timing in the Info&Bonus arm:

0. The bureaucracy announces the monitoring and incentive scheme. Nature draws realizations of C^T and C^H and reveals both of these costs to the teacher and the head-teacher.

1. The head-teacher chooses whether to make a side-contract offer to the teacher. A side-contract $< R(a), \tau >$ commits the head-teacher to monitor, specifies the report r that the head-teacher will send to all players following the action a and specifies the side-transfer τ that the teacher will pay to the head-teacher in the event that $r = 1$.

If the side-contract is accepted at Stage 1, the game continues as follows.

2. The teacher chooses whether to attend school, $a \in \{0, 1\}$.

3. The head-teacher monitors and sends the public report $r = R(a)$. If $r = 1$, the bureaucracy pays β to the teacher who then transfers τ to the head-teacher.

4. The bureaucracy chooses a policy decision $p \in \{0, 1\}$. Payoffs are realized and the game ends.

If the side-contract is not accepted at Stage 1, the game proceeds as in the Info arm except that at Stage 3 the bureaucracy transfers β to the teacher in the event that $r = 1$.

2.2 First Best

To provide a benchmark for later analysis, we begin by considering the outcome when both the teacher's cost C^T and her action a are public information. We continue to restrict the set of mechanisms available to the bureaucracy to simple bonus contracts for the teacher. Since the bureaucracy observes a there will never be a policy mistake, implying that we can ignore κ.

In the status quo with no bonus scheme, the teacher will attend school iff $C^T \leq \delta$, in which case $U^G = n\,\varepsilon^P$ (and otherwise $U^G = 0$). Now suppose that the bureaucracy offers to pay the teacher a bonus β if and only if she attends school. The teacher will attend iff $C^T \leq \delta + \beta$, in which case $U^G = n\,\varepsilon^P - \beta$ (and otherwise $U^G = 0$). It follows that the bureaucracy will adopt the bonus scheme iff the teacher's cost takes an intermediate value: $\delta < C^T < n\,\varepsilon^P + \delta$. Specifically, for these parameter values, the bureaucracy commits to pay a bonus of $\beta = C^T - \delta$ iff $a = 1$ (and otherwise nothing). Hence, in equilibrium, the teacher attends iff $C^T < n\,\varepsilon^P + \delta$.

2.3 Analysis

We now turn to our main theoretical analysis of the model in Section 2.1. Readers who wish to skip these derivations can proceed directly to Section 2.4 where we provide a graphical representation of the four testable predictions. Expected welfare comparisons are set out in Section 2.5.

2.3.1 Control arm

The equilibrium outcome is straightforward to establish. The teacher cares only about her participation cost and so attends iff $C^T \leq 0$. The probability of teacher attendance is therefore

$$\Pr[a = 1|\text{Control}] = \Pr[C^T \leq 0] = \frac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T}. \tag{2}$$

The bureaucracy cares only about minimizing the probability of making a policy mistake. Given the symmetric loss function, the bureaucracy will set $p = 1$ iff it believes the teacher is more likely to be present than absent. Anticipating the teacher's strategy, the bureaucracy deduces that the probability of teacher attendance is

$$\Pr[a = 1|m = 0] = \frac{\Pr[m = 0, a = 1]}{\Pr[m = 0]} = \Pr[C^T \leq 0] = \frac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} > 0.5,$$

and so it chooses $p = 1$. The probability of a policy mistake is therefore

$$\Pr[p \neq a|\text{Control}] = \Pr[C^T > 0] = \frac{\overline{C}^T}{\overline{C}^T - \underline{C}^T}. \tag{3}$$

It follows that the bureaucracy's expected payoff is

$$\begin{aligned} \mathrm{E}[U^G_{\text{Control}}] &= n\,\varepsilon^P \cdot \Pr[a = 1|\text{Control}] - \kappa \cdot \Pr[p \neq a|\text{Control}] \\ &= n\,\varepsilon^P \cdot \frac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} - \kappa \cdot \frac{\overline{C}^T}{\overline{C}^T - \underline{C}^T}. \end{aligned} \tag{4}$$

2.3.2 Info Arm

The equilibrium outcome is still fairly straightforward to establish. There are four different cases.

- *Truthful reporting of teacher presence*, $m = 1, r = a = 1$.
  This outcome arises if $0 < C^T \leq \delta$ and $C^H \leq \varepsilon^H$. The head-teacher knows that if she commits to monitor and report truthfully then the teacher will attend school (because $C^T \leq \delta$) giving her a payoff of $\varepsilon^H - C^H$. This is a better outcome for the head-teacher than if she does not commit to monitor, since the teacher will not attend school (because $C^T > 0$) giving her a payoff of 0. This outcome also arises if $C^T \leq 0$ and $C^H \leq 0$. The head-teacher knows that monitoring has no impact on the teacher, as he will always attend. Since the head-teacher derives utility from monitoring she chooses to do so.

- *Truthful reporting of teacher absence*, $m = 1, r = a = 0$.
  This outcome arises if $C^T > \delta$ and $C^H < 0$. The head-teacher knows that monitoring has no impact

on the teacher, as he will never attend. Since the head-teacher derives utility from monitoring she chooses to do so.

- *No monitoring of teacher absence*, $m = 0, a = 0$.
  This outcome arises if $C^T > \delta$ and $C^H > 0$. The head-teacher knows that monitoring has no impact on the teacher, as he will never attend. The head-teacher therefore refrains from costly monitoring. This outcome also arises if $C^T > 0$ and $C^H > \varepsilon^H$. The head-teacher knows that if she does not commit to monitor, the teacher will not attend school (because $C^T > 0$) giving her a payoff of 0. If she commits to monitor, the highest payoff that she can achieve is $\varepsilon^H - C^H < 0$ and so she again refrains from costly monitoring.

- *No monitoring of teacher presence*, $m = 0, a = 1$.
  This outcome arises if $C^T \leq 0$ and $C^H > 0$. The head-teacher knows that monitoring has no impact on the teacher, as he will always attend. The head-teacher therefore refrains from costly monitoring.

The probability of teacher attendance is therefore

$$\Pr[a = 1|\text{Info}] = \Pr[C^T \leq 0] + \Pr[0 < C^T \leq \delta, C^H \leq \varepsilon^H] = \frac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} + \frac{\delta(-\underline{C}^H + \varepsilon^H)}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}, \tag{5}$$

and the probability of monitoring is

$$\Pr[m = 1|\text{Info}] = \Pr[C^H \leq 0] + \Pr[0 < C^T \leq \delta, C^H \leq \varepsilon^H] = \frac{-\underline{C}^H}{\overline{C}^H - \underline{C}^H} + \frac{\delta\varepsilon^H}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}. \tag{6}$$

In contrast, to the Control arm, the bureaucracy now reaches three information sets, and so we must consider how it forms beliefs at each of these. The first information set is $m = 0$. Anticipating the teacher and head-teacher's strategies, the bureaucracy deduces that the probability of teacher attendance at this information set is

$$\begin{aligned}\Pr[a = 1|m = 0] = \frac{\Pr[m = 0, a = 1]}{\Pr[m = 0]} &= \frac{\Pr[C^T \leq 0, C^H > 0]}{\Pr[C^T \leq 0, C^H > 0] + \Pr[C^T > 0, C^H > 0] - \Pr[0 < C^T \leq \delta, 0 < C^H \leq \varepsilon^H]} \\ &= \frac{-\underline{C}^T \overline{C}^H}{(\overline{C}^T \overline{C}^H - \delta\varepsilon^H) - \underline{C}^T \overline{C}^H} > \frac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} > 0.5,\end{aligned}$$

and so it chooses $p = 1$. The second information set is $r = 1$. Since the bureaucracy knows that the head-teacher reports truthfully, it chooses p=1. The third information set is $r = 0$. Again, since the bureaucrat knows that the head-teacher reports truthfully, it chooses p=0. The probability of a policy mistake is therefore

$$\begin{aligned}\Pr[p \neq a|\text{Info}] &= \Pr[C^T > 0] - \Pr[C^T > 0, C^H \leq 0] - \Pr[0 < C^T \leq \delta, 0 < C^H \leq \varepsilon^H] \\ &= \frac{\overline{C}^T}{\overline{C}^T - \underline{C}^T} - \frac{-\underline{C}^H \overline{C}^T + \varepsilon^H \delta}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}.\end{aligned} \tag{7}$$

It follows that the bureaucracy's expected payoff is

$$
\begin{aligned}
\mathrm{E}[U^G_{\text{Info}}] &= n\,\varepsilon^P \cdot \Pr[a=1|\text{Info}] - \kappa \cdot \Pr[p \neq a|\text{Info}] \\
&= n\,\varepsilon^P \cdot \left(\frac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} + \frac{\delta(-\underline{C}^H + \varepsilon^H)}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)} \right) - \kappa \cdot \left(\frac{\delta(-\underline{C}^H + \varepsilon^H) + (\overline{C}^T - \delta) - \underline{C}^H)}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)} \right).
\end{aligned} \tag{8}
$$

2.3.3 Info&Bonus arm

We show in Appendix A that, under the parameter assumptions in (1), if the bureaucracy chooses the Info&Bonus arm, then the optimal β must exceed ε^H. This again gives rise to four cases, although crucially one equilibrium outcomes differs to the Info arm; there is now untruthful reporting.

- *Truthful reporting of teacher presence*, $m = 1, r = a = 1$.
  This outcome arises if $0 < C^T \leq \varepsilon^H$ and $C^H \leq \varepsilon^H + \beta - C^T$. There are two sub-cases. If $C^T \leq \delta$ and $C^H \leq \varepsilon^H$, the teacher's outside-option is a payoff of $-C^T$ (since the head-teacher will commit to monitor and report truthfully even if she cannot extract any incentive payment from the teacher). It follows that a side-contract of truthful reporting and a transfer of $\tau = \beta$ will be accepted, giving the head-teacher a payoff of $\varepsilon^H + \beta - C^H > 0$. This is better than the head-teacher can achieve via any other contract. Intuitively, information alone is enough to incentivise the teacher to attend, so, since the head-teacher has the bargaining power, she captures all of the bonus payment. If $\varepsilon^H \geq C^T > \delta$ and/or $\varepsilon^H + \beta \geq C^H > \varepsilon$, the teacher's outside-option is a payoff of 0 (since the head-teacher will not commit to monitor and report truthfully). It follows that a side-contract of truthful reporting and a transfer of $\tau = \beta - C^T$ will be accepted, giving the head-teacher a payoff of $\varepsilon + \beta - C^T - C^H > 0$. This is better than the head-teacher can achieve via any other contract. Intuitively, the head-teacher now has to share some of the bonus payment with the teacher to compensate him for his participation cost. This cost is sufficiently low relative to the head-teacher's private benefit from teacher presence to make such an 'incentive' side-contract worthwhile. This outcome also arises if $C^T \leq 0$ and $C^H \leq \beta$. The teacher's outside option is $-C^T > 0$. Hence a side-contract of a truthful report and a transfer of $\tau = \beta$ will be accepted, giving the head-teacher a payoff of $\varepsilon^H + \beta - C^H$. This is better than the head-teacher can achieve via no contract iff $C^H \leq \beta$. Since her monitoring cost is sufficiently low, the head-teacher offers a 'superfluous' side-contract simply to collect the bonus payment.

- *Collusion, untruthful reporting of teacher presence*, $m = 1, r = 1, a = 0$.
  This outcome arises if $C^T > \varepsilon^H$ and $C^H \leq \beta$. The teacher's outside option is 0. Hence a side-contract of untruthful report ($r = 1$ for any a) and a transfer of $\tau = \beta$ will be accepted, giving the head-teacher a payoff of $\beta - C^H > 0$. This is better than the head-teacher can achieve via any other contract. Intuitively, the head-teacher no longer finds it worthwhile to incentivise the teacher to attend (his participation cost is too high). Since her monitoring cost is sufficiently low, the head-teacher offers a 'collusive' side-contract simply to collect the bonus payment.

- *No monitoring of teacher absence*, $m = 0, a = 0$.
  This outcome arises if $C^H > \beta$ and $C^H > \varepsilon^H + \beta - C^T$. The head-teacher's monitoring cost is now so high that neither the incentive nor the collusive side-contact is worthwhile.

- *No monitoring of teacher presence*, $m = 0, a = 1$.
  This outcome arises if $C^T \leq 0$ and $C^H > \beta$. The monitoring cost is now too high, and so the head-teacher withdraws the 'superfluous' side-contract.

The probability of teacher attendance is therefore

$$\begin{aligned}\Pr[a = 1|\text{Bonus}] &= \Pr[C^T \leq 0] + \Pr[0 < C^T \leq \varepsilon^H, C^H \leq \varepsilon^H + \beta - C^T] \\ &= \tfrac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} + \tfrac{\varepsilon^H(-\underline{C}^H + \beta) + (\varepsilon^H)2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)},\end{aligned} \tag{9}$$

and the probability of monitoring (equivalently the probability of a bonus payment) is

$$\begin{aligned}\Pr[m = 1|\text{Bonus}] &= \Pr[C^H \leq \beta] + \Pr[\beta < C^H < \varepsilon^H + \beta - C^T] \\ &= \tfrac{\beta - \underline{C}^H}{\overline{C}^H - \underline{C}^H} + \tfrac{(\varepsilon^H)^2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}.\end{aligned} \tag{10}$$

The bureaucracy now reaches just two information sets. The first information set is $m = 0$. Anticipating teacher and head-teacher strategies, the bureaucracy deduces that the probability of teacher attendance at this information set is

$$\begin{aligned}\Pr[a = 1|m = 0] = \frac{\Pr[m = 0, a = 1]}{\Pr[m = 0]} &= \tfrac{\Pr[C^T \leq 0, C^H > \beta]}{\Pr[C^T \leq 0, C^H > \beta] + \Pr[C^T \geq > 0, C^T > 0, C^H > \beta] - \Pr[\beta < C^H < \varepsilon^H + \beta - C^T]} \\ &= \tfrac{-\underline{C}^T(\overline{C}^H - \beta)}{(\overline{C}^T(\overline{C}^H - \beta) - (\varepsilon^H)^2/2) - \underline{C}^T(\overline{C}^H - \beta)} > \tfrac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} > 0.5,\end{aligned}$$

and so it chooses $p = 1$. The second information set is $m = 1, r = 1$. Anticipating teacher and monitor strategies (in particular that the head-teacher may now report *untruthfully*), the bureaucracy deduces that the probability of teacher attendance at this information set is

$$\begin{aligned}\Pr[a = 1|m = 1, r = 1] = \frac{\Pr[m = 1, r = 1, a = 1]}{\Pr[m = 1, r = 1]} &= \tfrac{\Pr[C^T \leq \varepsilon^H, C^H \leq \beta] + \Pr[\beta < C^H < \varepsilon^H + \beta - C^T]}{\Pr[C^T \leq \varepsilon^H, C^H \leq \beta] + \Pr[\beta < C^H < \varepsilon^H + \beta - C^T] + \Pr[C^T > \varepsilon^H, C^H \leq \beta]} \\ &= \tfrac{(-\underline{C}^T + \varepsilon^H)(\beta - \underline{C}^H) + (\varepsilon^H)^2/2}{(-\underline{C}^T + \varepsilon^H)(\beta - \underline{C}^H) + (\varepsilon^H)^2/2 + (\overline{C}^T - \varepsilon^H)(\beta - \underline{C}^H)} > \tfrac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} > 0.5,\end{aligned}$$

and so again it chooses p=1. The probability of a policy mistake is therefore

$$\begin{aligned}\Pr[p \neq a|\text{Bonus}] &= \Pr[C^T > 0] - \Pr[0 < C^T \leq \varepsilon^H, C^H \leq \varepsilon^H + \beta - C^T] \\ &= \tfrac{\overline{C}^T}{\overline{C}^T - \underline{C}^T} - \tfrac{\varepsilon^H(-\underline{C}^H + \beta) + (\varepsilon^H)^2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}.\end{aligned} \tag{11}$$

It follows that the bureaucracy's expected payoff is

$$\begin{aligned}\mathrm{E}[U^G_{\text{Bonus}}] &= n\,\varepsilon^P \cdot \Pr[a = 1|\text{Bonus}] - \kappa \cdot \Pr[p \neq a|\text{Bonus}] - \beta \cdot \Pr[m = 1|\text{Bonus}] \\ &= n\,\varepsilon^P \cdot \left(\tfrac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} + \tfrac{\varepsilon^H(-\underline{C}^H + \beta) + (\varepsilon^H)^2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}\right) - \kappa \cdot \left(\tfrac{\overline{C}^T}{\overline{C}^T - \underline{C}^T} - \tfrac{\varepsilon^H(-\underline{C}^H + \beta) + (\varepsilon^H)^2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}\right) \\ &\quad - \beta \cdot \left(\tfrac{\beta - \underline{C}^H}{\overline{C}^H - \underline{C}^H} + \tfrac{(\varepsilon^H)^2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}\right).\end{aligned} \tag{12}$$

2.4 Testable Predictions

The analysis in Section 2.3 deliver a series of testable predictions. Prediction 1 follows from a comparison of (2), (5), and (9), Predictions 2 and 3 from a comparison of (3), (7), and (11), and Prediction 4 directly from (10) and the observation that, if the head-teacher monitors, then she marks the teacher present.

Prediction 1. *The probability of teacher attendance is (i) higher in the Info arm than in the Control arm, and (ii) higher in the Info&Bonus arm than in the Info arm.*

Prediction 2. *The probability of a policy mistake because the head-teacher does not monitor and the teacher is absent is (i) lower in the Info arm than in the Control arm, and (ii) lower in the Info&Bonus arm than in the Info arm.*

Prediction 3. *The probability of a policy mistake because the head-teacher falsely reports the teacher to be present is (i) equal in the Info and Control arms, and (ii) higher in the Info&Bonus arm than in the Info arm.*

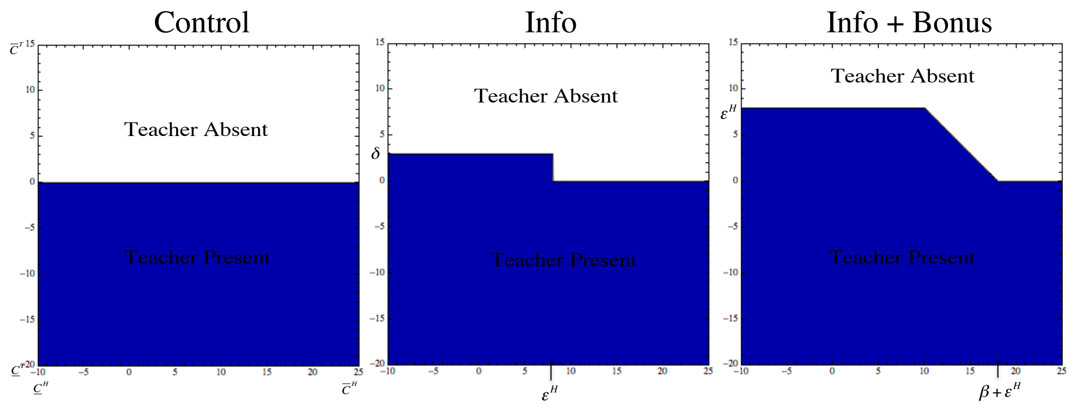
Prediction 4. *The probability of reported teacher attendance is (i) higher in the Info arm than in the Control arm, and (ii) higher in the Info&Bonus arm than in the Info arm. Hence, (iii) the probability of a bonus payment is positive in the Info&Bonus arm.*

The teacher present whenever $C^T < 0$; this is the baseline presence rate. Prediction 1 (i) can be seen in the first two panels of Figure 1. There are parameter values ($0 < C^T \leq \delta$ and $0 < C^H \leq \varepsilon^H$) where the possibility of reporting changes the teacher's behaviour from absence in the Control arm to presence in the Info arm. We refer to this as the *attendance effect* of local monitoring. The size of this effect is driven by the reputational cost δ and the head-teacher's valuation of teacher attendance ε^H. Prediction 1 (ii) can be seen in the final two panels of Figure 1. There are parameter values ($\delta < C^T \leq \varepsilon^H$ and $\varepsilon^H < C^H \leq \beta + \varepsilon^H - C^T$) where the additional possibility of a cash bonus changes the teacher's behaviour from absence in the Info arm to presence in the Info&Bonus arm. Hence, the attendance effect of local monitoring *and* incentives is greater than the attendance effect of local monitoring alone. The difference is increasing in ε^H and β, since it is these parameters that determine the size of the pie that is available to cover the bargaining parties' participation costs C^T and C^H.[7]

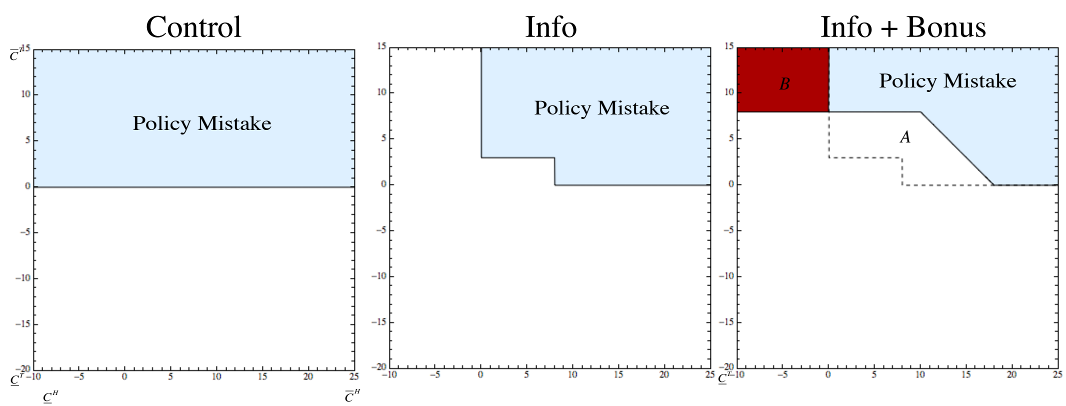
Prediction 2 (i) is driven by two forces. The first is simply the attendance effect of local monitoring. The possibility of reporting changes behaviour from no monitoring of teacher absence in the Control arm to truthful reporting of teacher presence in the Info arm. This eliminates the policy mistake that occurs in the Control arm, not by changing the policy which remains $p = 1$ but by changing a from 0 to 1. Second, there are parameter values ($C^T > \delta$ and $C^H < 0$) where the possibility of reporting has no impact on the teacher but changes the head-teacher's behaviour from no monitoring in the Control arm to truthful reporting in the Info arm. This eliminates the policy mistake that occurs in the Control arm, not by changing the teacher's action which remains $a = 0$ but by changing p from 1 to 0. This second force stems from direct utility gains from monitoring, and so its magnitude depends on $\underline{C}^H < 0$. Prediction 2 (ii) is driven by the difference in attendance effects. Region A in Figure 1(b) depicts parameter values where

[7]Obviously, the parameter restriction in (1), specifically $\beta > \varepsilon^H > \delta$, and the fact that β is a source of transferable utility are important here.

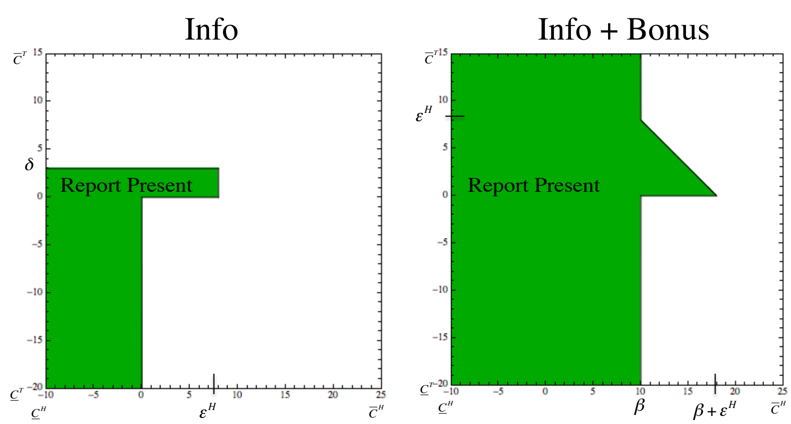
Figure 1: Teacher and monitor preferences and outcomes under alternative treatments



(a) Probability of teacher attendance (Prediction 1)



(b) Probability of a policy mistake (Prediction 2 and 3)



(c) Probability of reported teacher attendance (Prediction 4) the additional possibility of a cash bonus changes behaviour from no monitoring of teacher absence in the Info arm to truthful reporting of teacher presence in the Info&Bonus arm. This eliminates the policy mistake that occurs in the Info&Bonus arm, again not by changing the policy which remains $p = 1$ but by changing a from 0 to 1.

Prediction 3 (i) has a straightforward explanation: with no transferable utility on the table, the head-teacher either cannot, or has no incentive to, monitor and send a false report. Hence, the probability of a policy mistake due to false reporting is equal (to zero) in both the Control and the Info arms. Prediction 3 (ii) is also simple. Region B in Figure 1(b) depicts parameter values where the additional possibility of a cash bonus has no impact on the teacher but changes the head-teacher's behaviour from truthful reporting in the Info arm to false reporting in the Info&Bonus arm. This introduces a policy mistake in the Info&Bonus arm, not by changing the teacher's action which remains $a = 0$ but by changing p from 0 to 1.[8] Such collusive behaviour occurs because, with $C^T > \varepsilon^H$, it is no longer worthwhile for the head-teacher to compensate the teacher for his participation cost; she prefers instead to lie and pocket the entire bonus payment.

Together Predictions 2 and 3 give us the overall probability of a policy mistake. As one might expect, this is predicted to be lower in the Info arm than in the Control arm. However, contrary to common intuition, the overall probability of a policy mistake is not predicted to be lower in the Info arm than in the Info&Bonus arm. The difference is ambiguous and depends on the relative magnitude of Regions A and B. Indeed it is perfectly possible (as shown in Figure 1(b)) for Region A to exceed Region B, and hence for the introduction of the cash bonus to *reduce* the likelihood of a policy mistake.

Prediction 4 (i) follows from the fact that the head-teacher reports truthfully in the Info arm if $C^H < 0$ and/or there is an attendance effect of local monitoring. Prediction 4 (ii) is driven by the larger attendance effect in the Info&Bonus arm, and the head-teacher's incentive to report the teacher present and trigger the bonus payment. The shaded region in the final panel shows the parameter values (low C^H) where the bureaucracy makes a bonus payment to the teacher, Prediction 4 (iii).

2.5 Welfare Comparisons

Our goal is to establish whether the bureaucracy's expected payoff is higher in the Control arm with no scheme, in the Info arm with a monitoring scheme, or in the Info&Bonus arm with a monitoring and incentive scheme. Subtracting (4) from (8), we have

$$\mathrm{E}[U^G_{\text{Info}}] - \mathrm{E}[U^G_{\text{Control}}] =$$
$$n\,\varepsilon^P \cdot \overbrace{(\Pr[a=1|\text{Info}] - \Pr[a=1|\text{Control}])}^{>0,\ \text{Prediction 1 (i)}} - \kappa \cdot \underbrace{(\Pr[p \neq a|\text{Info}] - \Pr[p \neq a|\text{Control}])}_{<0,\ \text{Prediction 2 (i)}} > 0. \quad (13)$$

The first term on the RHS is the parental gain from increased teacher attendance, Prediction 1 (i). The second term on the RHS is the bureaucracy's gain from fewer policy mistakes, Predictions 2 (i). Clearly,

[8] Note that that this occurs even though the bureaucracy anticipates false reporting; i.e. it is *ex ante* optimal for the (Bayesian) bureaucracy to set $p = 1$.

the bureaucracy prefers the Info arm to the Control. Subtracting (8) from (12), we have

$$\begin{aligned} \mathrm{E}[U^G_{\text{Bonus}}] - \mathrm{E}[U^G_{\text{Info}}] = n\,\varepsilon^P \cdot \overbrace{(\Pr[a=1|\text{Bonus}] - \Pr[a=1|\text{Info}])}^{>0,\text{ Prediction 1 (ii)}} \\ - \kappa \cdot \underbrace{(\Pr[p \neq a|\text{Bonus}] - \Pr[p \neq a|\text{Info}])}_{?,\text{ Prediction 2 (ii) vs. Prediction 3 (ii)}} - \beta \cdot \underbrace{(\Pr[m=r=1|\text{Bonus}])}_{>0,\text{ Prediction 4 (iii)}}. \end{aligned} \tag{14}$$

The first term on the RHS represents the parental gain from increased teacher attendance, Prediction 1 (ii). The second term is of ambiguous sign. Theoretically, the bureaucracy could gain from fewer, or lose from more, policy mistakes, Predictions 2 (ii) and 3 (ii). The final term is the loss from paying out the cash bonus, Prediction 4 (iii).

Our theoretical model gives structure to this welfare analysis but, ultimately, gives only a partial welfare ranking. For this reason, we now turn to the data collected via the field experiment to estimate the above probabilities. These estimated probabilities will serve as a reality check of the model by enabling us to test Predictions 1-4. Together with the experimental value of β, they will also enable us to complete this welfare ranking for specified values of $n\,\varepsilon^P$ and κ.

3 Experimental Design and Data Description

3.1 Intervention Description and Experimental Design

In order to test the theoretical predictions of the model, we designed two local monitoring schemes where head-teachers were required to submit daily reports of teacher attendance. The two schemes are identical, except that in one intervention the reports also triggered bonus payments. The interventions are implemented in a random sample of 45 primary schools in rural Uganda.

Working with World Vision, we trained head-teachers, assisted by the deputy, in the use of a platform which allows them to report teacher attendance on a mobile device. This information, combined with a unique identification number of the school, teacher and monitor, is sent to a central database in Makerere University. The platform can be added to any Java-enabled phone, but we provide one phone per school to be sure. In all the intervention schools we re-broadcast a summary of results of teacher attendance to school stakeholders on a monthly basis via SMS. However, in a random sub-set of these schools teachers would also receive a monthly bonus of UShs 60,000 (USD23, or 23 per cent of their monthly salary) if they were reported as present every week that month. We refer to these schools as belonging to the *Info&Bonus* arm and the schools where the reports were not combined with bonus payments as the *Info* arm. No monitoring took place in the control schools.

The study took place in six different districts in 85 rural government schools in Uganda.[9] Stratifying by district, we randomly assigned 40 schools to the control, 25 schools to the *Info* arm and 20 schools to the *Info&Bonus* arm. The intervention was implemented in September 2012 at the beginning of the third school term and lasted for a year until the end of the second school term in 2013.

[9]An additional 95 schools were also allocated to pilots of other monitoring schemes, which are not the focus of this paper

Table 1: Descriptive Statistics

	Observations	Mean	Std. Dev.	Min.	Max.
Teacher characteristics					
Teacher presence	978	0.72	0.45	0	1
Proportion of female teachers	978	0.38	0.49	0	1
Average Teacher Age	950	36.43	9.14	19	62
School Characteristics					
Teachers	85	11.51	4.28	5	23
Pupils enrolled per school	85	538.06	298.80	74	1611

3.2 Data Description

Our analysis draws from three sources of data: Our own independent spot-checks of teacher attendance, monitor reports of teacher attendance, and a school survey.

First, we conducted random spot-checks of teacher attendance, both before the intervention started and during every term that the intervention took place: July 2012, November 2012 (Term 3), April/May 2013 (Term 1), and August 2013 (Term 2). This data is therefore at the teacher-day level: each observation is a different spot-check for a different teacher. We then matched this data-set of teacher attendance with the monitor reports that took place for the same teacher on the same day. Finally, we conducted a school survey both before the intervention started (July 2012) and after it was completed (November 2013), which provides basic information about school and teacher characteristics.

Table 1 provides descriptive statistics of key school and teacher characteristics. In our sample of 85 schools, there are on average 11.5 teachers per school, leading to total sample of 978 teachers, who are predominantly male (62%), with ages that range between 19 and 62. Average pupil enrollment is 538 per school and ranges between 74 and $1,611$. Average teacher attendance during the baseline was 72%. This is in line with what has been observed in earlier studies on Uganda and other low income countries (Chaudhurry *et al* 2006; Barr *et al*, 2013)

Table 2 shows the baseline balance of select variables. Columns (1) to (3) show the mean values for the control and each treatment arm; columns (4) and (5) indicates the coefficients of regressing each outcome variable on the treatment dummies, controlling for district fixed effects and clustering the standard errors at the school level; and column (5) shows the difference between the two treatment arms. The sample is clearly balanced across the treatment arms for most characteristics. Most importantly, there is no statistically significant difference between treatment arms in teacher attendance rate and school enrollment figures, two key outcome variables for this paper. We observe slight imbalance in the number of teachers, but only when comparing *Info* schools with the *Info&Bonus* schools.

Table 2: Balance Statistics

	Control	Info	Bonus	Info v Control	Bonus v Control	Bonus v Info
Teacher Presence	0.76	0.68	0.70	-0.07	-0.06	0.02
	(0.43)	(0.47)	(0.46)	(0.06)	(0.06)	(0.07)
Female	0.38	0.34	0.41	-0.03	0.03	0.05
	(0.49)	(0.47)	(0.49)	(0.04)	(0.04)	(0.04)
Teacher Age	36.85	36.29	35.90	-0.59	-0.85	-0.26
	(9.20)	(9.08)	(9.09)	(0.84)	(0.82)	(0.99)
Pupil Enrollment	583.48	499.17	537.56	-105.00	-46.10	58.90
	(398.58)	(282.57)	(325.23)	(84.45)	(85.47)	(80.48)
Teachers per School	11.23	10.60	12.68	-0.91	1.48	2.39**
	(4.23)	(3.07)	(5.02)	(0.87)	(1.06)	(1.04)

Note: Columns (1) to (3) show the mean values of key baseline characteristics in the control and respective treatment arms. Columns (4) and (5) indicates the coefficients of regressing each outcome variable on the treatment dummies, controlling for district fixed effects. Columns (5) shows the difference in coefficient sizes between the two treatments. Standard errors are in parentheses and clustered at the school level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

3.3 Empirical Strategy

Our main specification is a simple cross-sectional comparison across all treatment arms and is estimated using the following equation:

$$Y_{i,s,t} = \sum_{t=1}^{3} \delta_t + \gamma_1 \left(\text{Info}\right)_s + \gamma_2 \left(\text{Info\&Bonus}\right)_s + \rho_d + \varepsilon_{i,s,t} \tag{15}$$

where $Y_{i,s,t}$ is the outcome for teacher i in school s in post-treatment time period t; δ_t are time dummies for each of the three rounds of post-treatment data collection; ρ_d is district fixed effects; $(\text{Info})_s$ and $(\text{Info\&Bonus})_s$ refer to the two treatment dummies; $\varepsilon_{i,s,t}$ is the error term clustered at the school level. Our preferred specification thus pools the treatment impact across the three post-treatment rounds of data collection, as there is no evidence of a change in the impact of the program over time.[10]

Furthermore, when our dependent variable is teacher presence we can also control for baseline data, since we conducted independent spot-checks prior to the intervention:

$$Y_{i,s,t} = \theta Y_{i,s,PRE} + \sum_{t=1}^{3} \delta_t + \gamma_1 \left(\text{Info}\right)_s + \gamma_2 \left(\text{Info\&Bonus}\right)_s + \rho_d + \varepsilon_{i,s,t} \tag{16}$$

where s, $Y_{i,s,PRE}$ is baseline attendance for teacher i in school s.

Finally, when we only compare the two treatment arms we estimate the following equation from a restricted sample which excludes the control schools:

$$Y_{i,s,t} = \sum_{t=1}^{3} \delta_t + \gamma_2 \left(\text{Info\&Bonus}\right)_s + \rho_d + \varepsilon_{i,s,t} \tag{17}$$

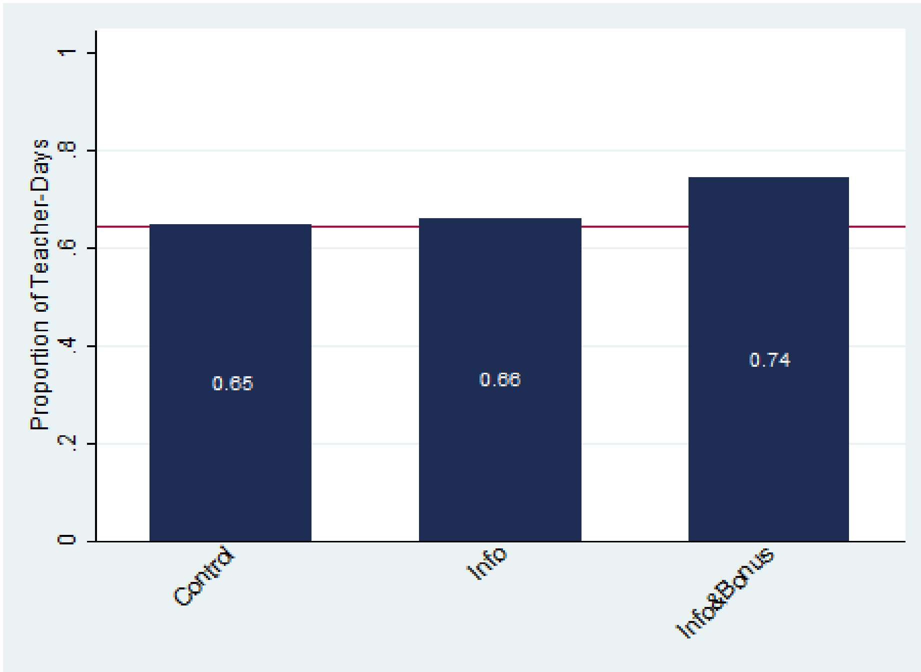
4 Results

In this section we test for the theoretical predictions of the model. In summary, we find that predictions 1, 2 and 4 are all confirmed by the data: teacher attendance is higher, non-reports of absent teachers less frequent, and reported presence higher in the *Info&Bonus* schools compared to both *Info* and *Control* schools. Prediction 3 is not supported by the data, as head teachers are equally likely to misreport teacher presence in both treatment arms. To empirically assess policy effectiveness, we estimate total policy mistakes and find that the bureaucrat would commit fewer mistakes under the *Info&Bonus* arm, compared to *Info*. Increasing the financial stakes of local monitoring is thus effective at both increasing teacher attendance and reducing policy mistakes.

Figures 2, 3 and 4 present the empirical tests for the theoretical predictions, as they were outlined in Figure 1a, b, and c of Section 2. Although it is easier to present our results graphically, Table 3 shows the regression results that support these graphs. Columns (1) and (3) to (6) report the results of a simple cross-sectional comparison, estimated using equations 15 and 17. Column (2) also controls for baseline presence and is thus estimated using equation 16. Panel (b) shows tests for the theoretical predictions of the model. The first row in panel (b) is the coefficient on the (Info) treatment dummy. The second

[10] In the appendix we also allow for time-variant treatment effects.

Figure 2: Teacher Presence



Note: Figure is based on 4493 teacher-days with independent spot-checks.

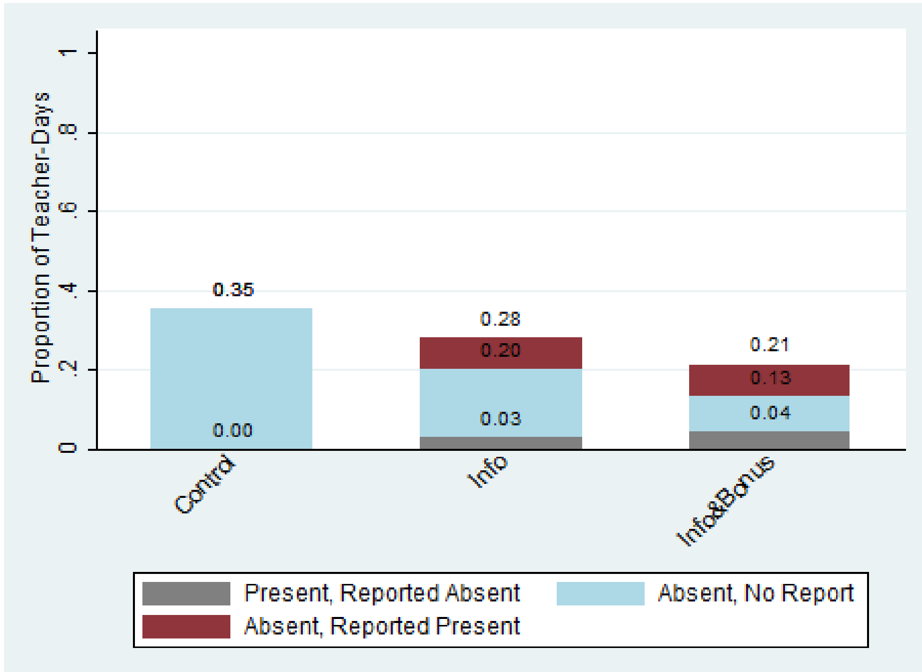
row in panel (b) is the coefficient on (Info&Bonus) from equation 17 and indicates the mean difference between *Info* and *Info&Bonus* schools.

We see from Figure 2 that teacher attendance increases when a bonus is introduced. On the days that we conducted independent, spot-checks teachers are 8 and 9 percentage points more likely to be present in the *Info&Bonus* schools compared to both the *Info* and *Control* schools. Column (1) shows that this difference is statistically significant, although attendance is not significantly higher in the *Info* schools relative to *Control*. Prediction 1(i) is thus supported by the data.

Next we show that policy mistakes due to non-reports of an absent teacher is lower when under financial incentives. To assess the likelihood of committing policy mistakes, we match independent spot-check of teacher attendance with the monitor report for the same teacher on the same day. Figure 3 reports a simple comparison across treatments, with light-blue bars reflecting the proportion of total teacher-days that the teacher was absent and no report was submitted. This number is 18 percentage points higher in the *Control* than in the *Info* arm (35% rather than 17%); and 8 percentage points higher in the *Info* than the *Info&Bonus* arm (17% rather than 9%). Column (3) in table 3 shows that these differences are statistically significant at the 1% and 10% levels respectively. Prediction 2(i) and 2(ii) are thus also supported by the data.

The maroon bars in Figure 3 reflect the proportion of teacher-days where an absent teacher is falsely reported as present. Contrary to the theoretical predictions, we also observe false reporting in the Info arm. Furthermore, the headteacher is no more likely to falsely report an absent teacher as present in the *Info&Bonus* arm, confirmed by the estimation results in column (4) in table 3. Prediction 3 is thus not supported by the data. For completeness, Figure 3 included the grey bars which reflect teacher-days

Figure 3: Policy Mistakes



Note: Figure is based on 4493 teacher-days with independent spot-checks.

where the monitor falsely reported a present teacher as absent. In both treatment arms the number is close to zero, and we consider this to be reporting error.

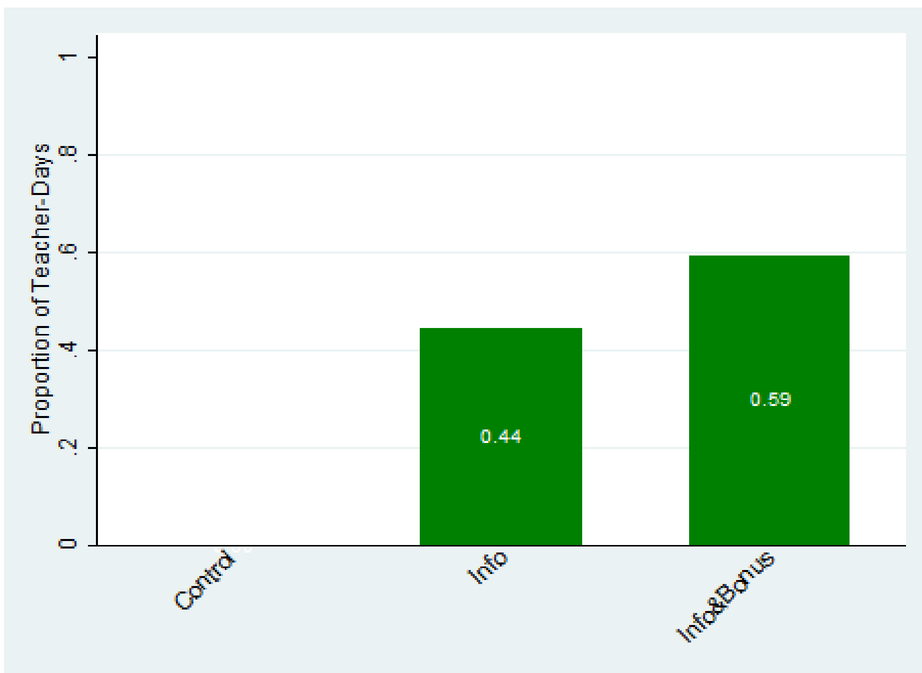
Theory was ambiguous as to which of the above two policy mistakes would dominate, but we can estimate this empirically. The size of the bars in figure 3 indicate the total sum of all policy mistakes. The policymaker turns out to be close to 7 percentage points more likely to commit a policy mistake in the *Info* than the *Info&Bonus* schools, and this is significant at the 10% level, as shown in column (5) of table 3.[11]

Finally, figure 4 indicates the frequency of a monitoring reporting a teacher as present.[12] In the *Info&Bonus* schools, teachers were reported as present in 59% of teacher-days. Reported presence was about 15 percentage points lower in the *Info* arm. Prediction (4) is therefore also supported by the data.

[11]Including the negligible proportion of teacher-days where a present teacher was falsely reported as absent.

[12]We use the restricted sample of teacher-days that we conducted independent spot-checks to assure comparability and consistency across the empirical tests. Results hold for the larger sample of all monitor reports.

Figure 4: Reported Teacher Attendance



Note: Figure is based on 4493 teacher-days with independent spot-checks.

Table 3: Results

Panel (a): *Regression Analysis*	(1)	(2)	(3)	(4)	(5)	(6)
	Teacher Presence		Abs, No Rep	Abs, Rep Present	Mistake	Rep Present
Info	0.0105	0.0145	-0.180***		-0.0724**	
	(0.04)	(0.04)	(0.04)		(0.04)	
Info&Bonus	0.0919**	0.0934***	-0.264***	0.00382	-0.139***	0.145**
	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)	(0.07)
Strata indicators	Yes	Yes	Yes	Yes	Yes	Yes
Baseline Control	No	Yes	No	No	No	No
Panel (b): *Test Theoretical Predictions*						
	Prediction 1		Prediction 2	Prediction 3		Prediction 4
(i): Info - Control	0.0105	0.0145	-0.180***		-0.0724**	
	(0.04)	(0.04)	(0.04)		(0.04)	
(ii): Info&Bonus - Info	0.0814*	0.0788*	-0.0843**	0.00382	-0.0669*	0.145**
	(0.04)	(0.04)	(0.04)	(0.02)	(0.04)	(0.07)
Obs	4493	4493	4493	2432	4493	2432
Control mean	0.647	0.647	0.353		0.353	

Note: The data is based on independent spot-checks of teacher attendance, matched with monitor reports that were submitted for the same teacher on the same day. Column (1), and columns (3) to (6) report regression results on post-treatment data only, estimated using equation 15 and 17. Column (2) controls for baseline attendance, based on equation 15. In columns (1) and (2) the independent variable is teacher attendance; in column (3) it is teacher-days where a teacher was absent but no report was submitted; in column (4) it is teacher-days that the monitor falsely reported an absent teacher as present; in column (5) it is the total number policy mistakes, including both false reports and teacher-days where a teacher was present but no report submitted; in column (6) it is reported presence. The final four rows report results as they relate to the theoretical predictions of the model. Row (i) reports coefficient on (Info) treatment dummy; row (ii) reports the difference in coefficient sizes between the two treatment dummies; Standard errors are in parentheses and clustered at the school level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

To summarize, predictions 1, 2 and 4 are all supported by the data. Contrary to our theory, we find that headteachers in the *Info* schools also falsely report absent teachers as present and are no more likely to do so when reports trigger bonus payments.[13] Furthermore, whereas the theoretical prediction was ambiguous as to which policy mistakes from predictions 2 and 3 dominate— no report versus false report on days when the teacher is absent— we can test this empirically and find that policy mistakes are *more* likely when no financial incentives are attached to the reports.

4.1 Robustness

As a robustness check we test for a Hawthorne effect. The reader might be concerned that our own independent spot-checks lead to higher teacher attendance. This could bias our results if the impact of our presence varies between treatment arms. For example, teachers in the *Info&Bonus* arms might be more responsive to our visits, if they believe it could have an implication for their bonus payments. To test for a Hawthorne effect, we visited some schools more than once each round of spot-checks and randomly varied the frequency of visits: some schools received three visits, and some schools only received one visit. We regress the number of spot-check visits on teacher attendance, controlling district and period fixed effects (not shown). We find no evidence of a Hawthorne effect, since the number visits does not significantly impact teacher attendance.

5 Welfare analysis

The preceding section has taken empirically testable propositions from the theoretical framework of Section 2 and shown these to be broadly confirmed by our experimental results. This *positive* result does not answer the welfare question: what *should* a policymaker do? That answer depends not only on the magnitude of behavioral responses to the intervention, but also on the social values placed on these outcomes. To address this question, we use the experimental results to evaluate welfare using the government bureaucracy's expected payoff as the relevant criterion, as proposed in Section 2.[14]

We focus here on what we consider to be the more interesting welfare comparison: moving from an unincentivized to an incentivized local monitoring scheme. Plugging the results of the experiment back into the welfare criterion in equation 14, an incentive scheme is only welfare-enhancing if the cost of bonus payments (in our case, USD$2,250$ per school) is less than both the social benefit from the increased attendance, $0.08 \cdot n\varepsilon^P$, and the reduction in policy mistakes due to the observed improvement in the quality of information, $0:07 \cdot \kappa$. [15]

To summarize, for a broad set of parameter assumptions we find that it is welfare-enhancing to pay for locally monitored teacher attendance. In fact, for most parameter values the net gain from attaching bonus payments to local reports is positive even without considering the cost of policy mistakes, κ.

[13] This suggests that there is some degree of transferable utility in teachers' reputational cost, δ, of being reported as present.

[14] We describe this as broad confirmation because the model did not predict the finding that the head-teacher would falsely report the teacher to be present in the Info arm or falsely report the teacher to be absent under either treatment arm. We account for this discrepancy in what follows because we focus on the overall probability of a policy mistake.

[15] The welfare comparison between *Info* and *Control* is relatively simple. Plugging the estimates in the first row of Table 3 Panel B columns (2) and (5) into equation (13), we get $\mathrm{E}[U^G_{\text{Info}}] - \mathrm{E}[U^G_{\text{Control}}] = n\varepsilon^P \cdot 0.01 + \kappa \cdot .07 > 0$.

Since the results of the experiment indicate that the quality of information actually *improved* due to the introduction of bonus payments, it is reasonable to conclude that welfare improved, even under some extremely conservative assumptions.

5.1 Enrollment Gains

As a basis for our welfare comparisons, we observe that enrollment expanded dramatically in the *Info&Bonus* schools, relative to both *Info* and Control schools. Table 4 reports the impact of the intervention on enrollment, using two different sources of data. The results in column (1) rely on our survey data and follows the same estimation strategy of equation 15, with the one exception that there is now only one post-intervention period.[16] Schools in the *Info&Bonus* arm report on average 47 more students enrolled compared with *Control*, and 70 pupils compared to *Info* schools. This represents a sizable increase of roughly 8.4 and 13 per cent over the average enrollment in the *Control* and *Info* arms respectively. Figure 5 displays these results graphically and shows that, for each grade, average enrollment is higher in the *Info&Bonus* than the *Info* schools.

This finding is corroborated in column (2), which reports school enrollment of a cohort of 20 pupils were surveyed in 2010 as part of a previous study. At the time these pupils were selected as representative of those enrolled in Primary 3. We tracked the enrollment outcomes of these children during our post-treatment school survey in November 2013. In the control schools, only 34% of these children are still enrolled in the same school, while pupils in *Info&Bonus* schools are 13.8 and 9 percentage points— or 40 and 23 percent— more likely to still be enrolled relative to *Control* and *Info* schools respectively.

The strikingly similar results across two different data sets allows us to conclude that enrollment increased due to the introduction of financial incentives. However, higher enrollment does not necessary imply that pupils are progressing to higher grades. In order to estimate the welfare impact, we need to model how the higher enrollment equates to higher attainment. We turn to this below.

5.2 Financial Returns to Schooling

To estimate the welfare gains attributable to enrollment decisions, we proceed in three steps. First, we model the impacts of the increase in enrolment on cumulative years of schooling. In particular, we back out the portion of the enrollment gain that is due to "averted drop-outs",[17] rather than grade repeaters, and then conservatively assume that each averted drop-out will progress only one additional grade before dropping out. Second, we use data from the 2011/12 Uganda National Panel Survey (UNPS) to estimate the net present value (NPV) of future lifetime earnings, for each level of grade attainment. Third, we calculate the increase in NPV for the pupils who gain a year of schooling, given different assumptions of the causal impact of schooling on earnings.[18]

[16] Survey data is missing in two schools for 2012, due to enumerator error. In these schools we imputed the EMIS 2012 data. We feel confident doing this, because our enrollment figures correspond closely to the EMIS data. In fact, the 2013 figures were exactly the same for the two schools with absent 2012 data. Results hold when we drop those two schools from the sample.

[17] School dropouts is a serious concern in Uganda, where only 30% of pupil who enroll in grade one make it to grade 7 (Ministry of Education and Sports, 2014:121). Similarly, in the control schools in our study, the number of pupils in grade 7 is on average only 40% of the number of grade 1 pupils.

[18] We are again being conservative by ignoring the non-financial benefits to education

Table 4: Enrollment impacts

	(1) Enrollment	(2) Tracked Cohort
Info	-23.52	0.0469
	(26.15)	(0.06)
Info&Bonus	46.50*	0.138**
	(24.12)	(0.05)
Strata indicators	Yes	Yes
Baseline Control	Yes	No
Obs	85	1140
Info=Info&Bonus: p-value	0.016	0.163
Control=Info&Bonus: p-value	0.058	0.014
Control mean	556.300	0.344

Note: Each column reports a separate regression, using the specification of equations 16 and 15. Dependent variables are (1) total school enrollment; and (2) a binary indicator of enrollment at follow-up for a sample of pupils tracked from the pre-intervention period. p-values reported for test of the equality of the *Info* and *Info&Bonus* treatment arms. Standard errors are in parentheses and clustered at the school level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

Figure 5: Average Enrollment by Grade

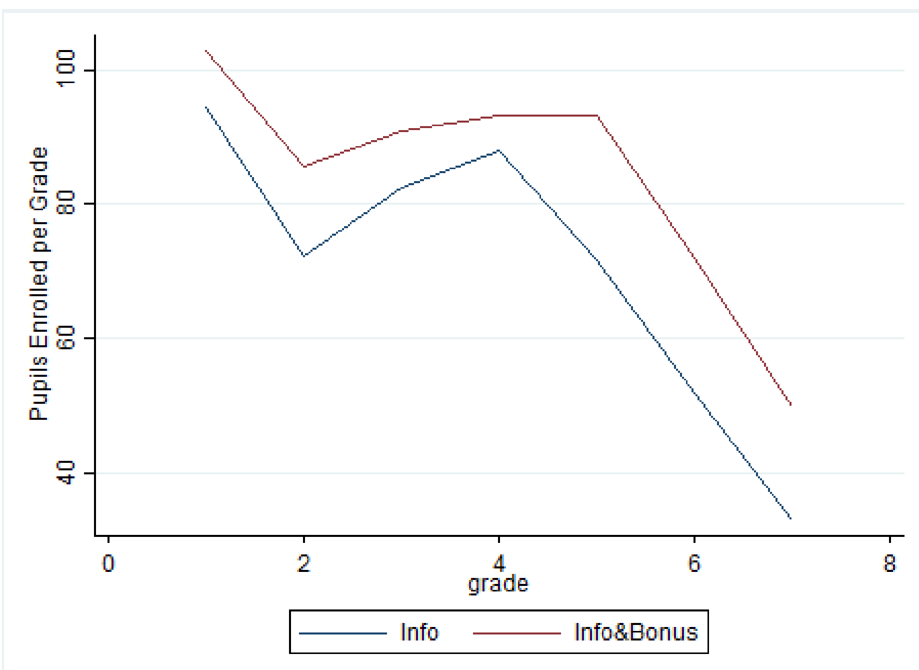


Table 5 shows different estimates for the total per-school gain in NPV of future life-time earnings, based on different parameter assumptions and data sources (see Section C in the Appendix for the calculations underlying these figures). We show results for two different discount rates,[19] two different values for the causal impacts of an additional year of education on wages,[20] and two different values of causal impact of an additional year of education on the probability of entering formal employment.[21] The results are surprisingly similar between the two different data sources, even though we used very different methods to estimate the gains in grade attainment.

Comparing the total financial benefit due to the improvement in teacher attendance with the per-school costs of bonus payments of USD 2,250, we see that for most parameter values it is welfare-enhancing to pay for locally monitored reports of teacher presence, even when not considering the reduction in policy mistakes, κ. However, for the most conservative parameter values (in red), we require some positive value of κ for the program to be welfare-enhancing. In these cases, it is then important to also consider the additional benefit of the program on the quality of information. Since even under these conservative assumptions the financial gain is very similar in magnitude to the cost of bonus payments, it is reasonable to conclude that the program is welfare-enhancing after also considering the benefits of improved quality of information.

Table 5: Per school difference between *Info&Bonus* with *Info* in the net present value of future lifetime earnings

		Administrative Data				**Cohort Data**			
Returns to educ. (wage)		**6.8%**		**15.9%**		**6.8%**		**15.9%**	
Returns to educ. (empl.)		**0.0%**	**0.8%**	**0.0%**	**0.8%**	**0.0%**	**0.8%**	**0.0%**	**0.8%**
Discount rate:	**12%**	1380	2589	3912	4827	1673	2138	3228	3985
	10%	2114	3294	4943	6122	1745	2722	4081	5057

Note:

To summarize, this exercise in cost-benefit analysis highlights the importance of a theoretical framework to make welfare calculations. Our theoretical framework provided us with a structure to incorporate both the value of information and value of teacher attendance in one welfare criteria. Furthermore, since the results of the theory demonstrated an actual *improvement* in the quality of information, this allowed us to compare the pure financial gains with costs of the bonus payments, as a sufficient condition for welfare improvement; and place a bound of κ of the minimum cost of policy mistakes required for welfare-improvement.

[19] A discount rate of 10% was used by both Ozier (2015) and Baird et al (2011) in their welfare analysis of a deworming program in Kenya. We compare this with a more conservative value of 12%.

[20] 6.8 is the most conservative value from studies that use quasi-experimental designs to credibly estimated the causal impact of education on earnings in developing countries (Card 2001, Duflo 2001). (Montenegro & Patrinos 2014) estimate a far larger value of 15.9% for Uganda, but their analysis is merely an ordinary-least-squared estimate on a cross-sectional data-set, so they do not adequately account for selection programs. The true value reasonable falls between these two estimates

[21] To the best of our knowledge, no studies in developing countries have looked at the impact of education on gaining formal employment. As a best guess, we rely on our own estimate used to calculate the NPV of future earnings (section C, in the Appendix). The most conservation alternative assumption is that education has no impact on the probability of entering the formal labor market.

5.3 Accounting for learning outcomes

Increases in years of completed schooling are not the only mechanism through which teacher attendance may impact pupil outcomes. Most obviously, teacher attendance may also have direct effects on learning outcomes, as has been shown elsewhere (Duflo et al. 2012). There is a danger of double-counting benefits: estimates of the return to additional years of education embody the learning gains that accompany these, and to our knowledge there do not exist experimental or quasi-experimental results that could cleanly separate the contributions of each. Nonetheless, as an alternative to bounding the required κ for policymakers to choose pay-for-performance, we consider valuing the impacts of learning gains. As we show below, the substantial impacts on enrollment imply that only quite wide bounds can be placed on any learning gains that may have been induced by these interventions.

Did the program lead to improved learning? To test for this we also conducted independent numeracy and literacy tests on a random sample of 40 pupils in each school, 20 in grades 3 and 6 respectively. However, the large impact of the treatment on pupil retention cautions against testing for learning outcomes: If the worst-performing pupils are most likely to drop out, then any impact on learning outcomes will be biased downwards due to changed composition of pupils between treatment arms.[22] To test for this, we conduct Lee Bounds on the learning outcomes. To calculate the lower bound of the impact, the Lee Bound estimator drops the best-performing pupils in the group with lower attrition rate (in our case, the *Info & Bonus* treatment arm) until the the shared of observations with observed outcomes is equal. Similarly for the upper bounds, the estimator drops all the worst-performing pupils.

Table 6 shows the results for constructing Lee Bounds on learning outcomes, based on the 57 schools where we also surveyed pupils before the intervention. The odd-numbered columns show results for the literacy test, the even-numbered columns the results for numeracy. The first two columns compare the difference in learning outcomes between the *Info&Bonus* arm and the Control schools; the next two columns compare *Info* and*Info&Bonus*; the final two columns compare *Info* and Control. The first and third row indicates the lower and upper bounds respectively. The third-last row shows the total number of observations for the full sample, including those who dropped out (i.e. 20 pupils per school), the second-last row shows the number of pupils still at the school; the final row shows the proportion of non-missing observations that were dropped from the treatment arm with higher retention. It is clear from table 6 that large different rates of retention across treatment arms invalidated any assessment of learning outcomes: Depending on the most extreme assumptions of sample selection, on could infer that the *Info&Bonus* has a statistically significantly positive or negative impact. Note that the range is largest in columns (1) and (2). This is because attrition was largest in the Control arm, so a larger proportion of observations in the *Info&Bonus* group need to be dropped.

6 Conclusions

This paper has provided theory and evidence on the impacts of alternative designs for local monitoring of teacher attendance in Uganda. Theory provides not only a positive set of testable predictions, but also a normative criterion for deciding amongst alternatives, which guides our analysis. Consistent with the

[22]It is, in fact, the case in our sample that those who dropped out performed worse in both numeracy and literacy tests in the baseline, although the difference is not statistically significant.

Table 6: Lee Bounds

	Info&Bonus vs Control		Info&Bonus vs Info		Info vs Control	
	(1)	(2)	(3)	(4)	(5)	(6)
	Literacy	Numeracy	Literacy	Numeracy	Literacy	Numeracy
lower	-7.519***	-5.860***	-3.890	-4.424	-4.607*	-2.623
	(2.02)	(1.67)	(2.89)	(2.83)	(2.70)	(2.19)
upper	9.637***	9.499***	8.936***	6.875**	0.436	1.982
	(3.09)	(2.91)	(3.15)	(2.89)	(2.50)	(2.14)
Total Obs	860	860	620	620	800	800
Selected Obs	282	282	224	224	236	236
Ratio	0.288	0.288	0.199	0.199	0.111	0.111

Note: Each column reports a separate regression using the Lee Bounds estimator. The odd-numbered columns report results from literacy test and the even-numbered columns report results from numeracy tests. Standard errors, clustered at the school level, are in parentheses. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

model, we find that teacher attendance, monitoring frequency, and the number of false reports all increased with the introduction of financial incentives. More surprisingly, but again consistent with the theory, we find that the number of policy mistakes actually decreased: there were more false reports but this effect was counter-balanced by more reports in general and (hence) fewer mistakes due to a lack of information. Taken together, our results suggest that social welfare was higher with financial incentives. This outcome results from a combination of low monitoring costs and high alignment of monitor preferences with government objectives. Such preferences and costs are key parameters determining the likely success of delegated monitoring schemes applied to other contexts.

References

Banerjee, A., Duflo, E. & Glennerster, R. (2007), 'Putting a band-aid on a corpse: incentives for nurses in the Indian public health care system', *Journal of the European Economic Association* **6**, 487–500.

Bénabou, R. & Tirole, J. (2003), 'Intrinsic and extrinsic motivation', *Review of Economic Studies* **70**, 489–520.

Bénabou, R. & Tirole, J. (2006), 'Incentives and prosocial behavior', *American Economic Review* **96**(5), 1652–1678.

Card, D. (2001), 'Estimating the return to schooling: Progress on some persistent econometric problems', *Econometrica* **69**(5), 1127–1160.

Chaudhury, N., Hammer, J., Kremer, M., Muralidharan, K. & Rogers, F. H. (2006), 'Missing in action: Teacher and health worker absence in developing countries', *Journal of Economic Perspectives* **20**(1), 91–116.

Deci, E. L. (1971), 'Effects of externally mediated rewards on intrinsic motivation', *Journal of Personality and Social Psychology* **18**, 105–115.

Duflo, E. (2001), 'Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment', *American Economic Review* **91**(4), 795–813.

Duflo, E., Greenstone, M., Pande, R. & Ryan, N. (forthcoming), 'Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India', *Quarterly Journal of Economics* .

Duflo, E., Hanna, R. & Ryan, S. P. (2012), 'Incentives work: Getting teachers to come to school', *American Economic Review* **102**(4), 1241–1278.

Glewwe, P., Ilias, N. & Kremer, M. (2010), 'Teacher incentives', *American Economic Journal: Applied Economics* **2**(3), 205–227.

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

Jacob, B. & Levitt, S. (2003), 'Rotten apples: An investigation of the prevalence and predictors of teacher cheating', *Quarterly Journal of Economics* **118**(3), 843–877.

Linden, L. L. & Shastry, G. K. (2012), 'Grain inflation: Identifying agent discretion in response to a conditional school nutrition problem', *Journal of Development Economics* **99**, 128–138.

Montenegro, C. & Patrinos, H. A. (2014), 'Comparable estimates of returns to schschool around the world', World Bank, Policy Research Working Paper Series 7020.

Muralidharan, K. & Sundararaman, V. (2011), 'Teacher performance pay: Experimental evidence from India', *Journal of Political Economy* **119**(1), 39–77.

VA Office of Inspector General (2014), 'Review of patient waiting times, and alleged patient deaths at the Phoenix health care system', Interim Report, Veterans Health Administration.

Vann, M. G. (2003), 'Of rats, rice, and race: The great Hanoi rat massacre, an episode in French colonial history', *French Colonial History* **4**, 191–203.

Wane, W. & Martin, G. H. (2013), 'Education and health services in Uganda: Data for results and accountability', World Bank, Service Delivery Indicators Launch Report, Washington DC.

Appendix A Further theoretical results

Here we justify our claim that under Assumption 1, if the bureaucracy chooses the Info+Bonus arm, then the optimal bonus level is $\beta > \varepsilon^H$. There are three cases to consider.

Case (i) $\beta \geq \varepsilon^H$.

This is the case discussed in the text. Given $\beta > \varepsilon^H$, the head-teacher is never constrained when making a side-contract offer to the teacher; i.e. if the head-teacher wants to incentivise the teacher to attend, then there is enough bonus to compensate the teacher for C^T. The bureaucracy's expected payoff is given in (12). Note that $\frac{\partial^2 \mathrm{E}[U^G_{\text{Bonus}}]}{\partial \beta^2} < 0$, implying that the objective is concave in β. The intuition here is that eventually (when $\beta >> \varepsilon^H$) there comes a point where it is not worthwhile increasing β. Doing so results in more collusion and socially wasteful monitoring for only a small gain in teacher attendance.

Case (ii) $\varepsilon^H > \beta \geq \delta$.

It is straightforward to show that, in this case, the objective is always increasing in β. It follows that the bureaucracy would never want to choose β in this region. This is because the head-teacher is constrained; she wants to offer an incentive side-contract but is unable to because she cannot cover the teacher's participation cost.

Case (iii) $\delta > \beta$.

In this case, the objective is always decreasing in β. It follows that the bureaucracy would never choose $\beta > 0$ in this region. This is because β yields only a small attendance gain relative to collusion.

All that remains is to establish whether the bureaucracy prefers $\beta \geq \varepsilon^H$ to $\beta = 0$. Note that

$$\mathrm{E}[U^G_{\text{Bonus}}]|_{\beta=\varepsilon^H} - \mathrm{E}[U^G_{\text{Bonus}}]|_{\beta=0} = (n\,\varepsilon^P + \kappa)\left(\frac{(\varepsilon^H-\delta)(-\underline{C}^H+\varepsilon^H)+(\varepsilon^H)^2/2)}{(\overline{C}^T-\underline{C}^T)(\overline{C}^H-\underline{C}^H)}\right) - \varepsilon^H\left(\frac{(\overline{C}^T-\underline{C}^T)(-\underline{C}^H+\varepsilon^H)+(\varepsilon^H)^2/2}{(\overline{C}^T-\underline{C}^T)(\overline{C}^H-\underline{C}^H)}\right).$$

The first term on the RHS is the gain from paying the bonus (higher attendance and fewer policy mistakes), while the second term is the loss from paying the bonus $\beta = \varepsilon^H$. Under Assumption 1,

$$n\,\varepsilon^P\left(\frac{(\varepsilon^H-\delta)(-\underline{C}^H+\varepsilon^H)+(\varepsilon^H)^2/2)}{(\overline{C}^T-\underline{C}^T)(\overline{C}^H-\underline{C}^H)}\right) - \varepsilon^H\left(\frac{(\overline{C}^T-\underline{C}^T)(-\underline{C}^H+\varepsilon^H)+(\varepsilon^H)^2/2}{(\overline{C}^T-\underline{C}^T)(\overline{C}^H-\underline{C}^H)}\right) > 0.$$

Hence we can focus on the case where $\beta > \varepsilon^H$.

Appendix B Additional Empirical Analysis

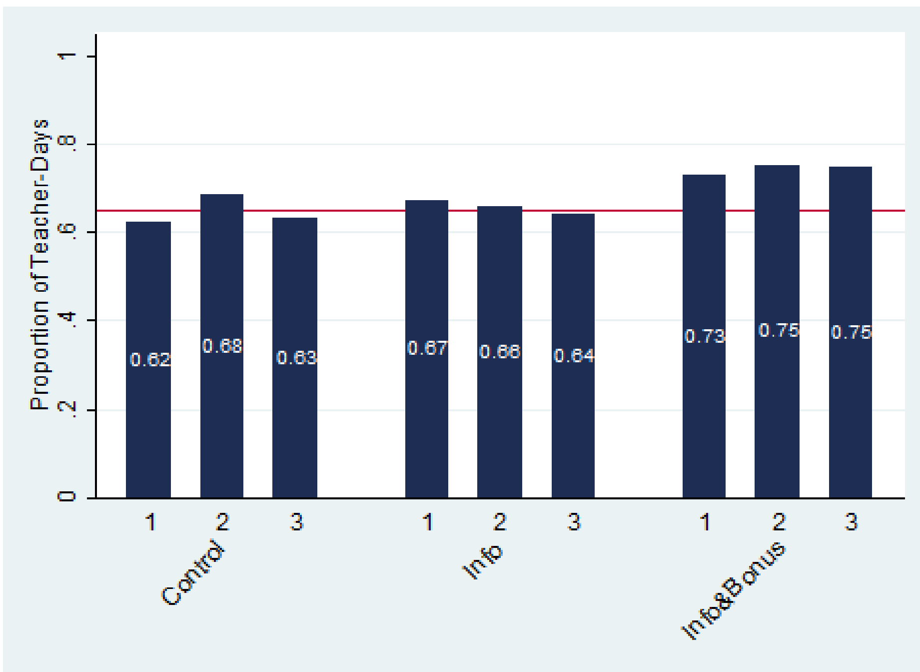
Appendix B.1 Time-varying treatment effects

In this appendix, we allow for a more general specification where the treatment effects vary across terms, in order to test for learning across terms. We estimate the following equation:

$$Y_{is,t} = \sum_{t=1}^{r} \delta_t + \gamma_{1t} \left(\text{info}_{is,t}\right) + \gamma_{2t} \left(\text{bonus}_{is,t}\right) + \theta Y_{is,PRE} + \rho_d + \varepsilon_{is,t} \tag{18}$$

where treatment effects γ_{1t} and γ_{2t} are allowed to vary across time.

Figure B.1: Teacher presence effects by term of exposure



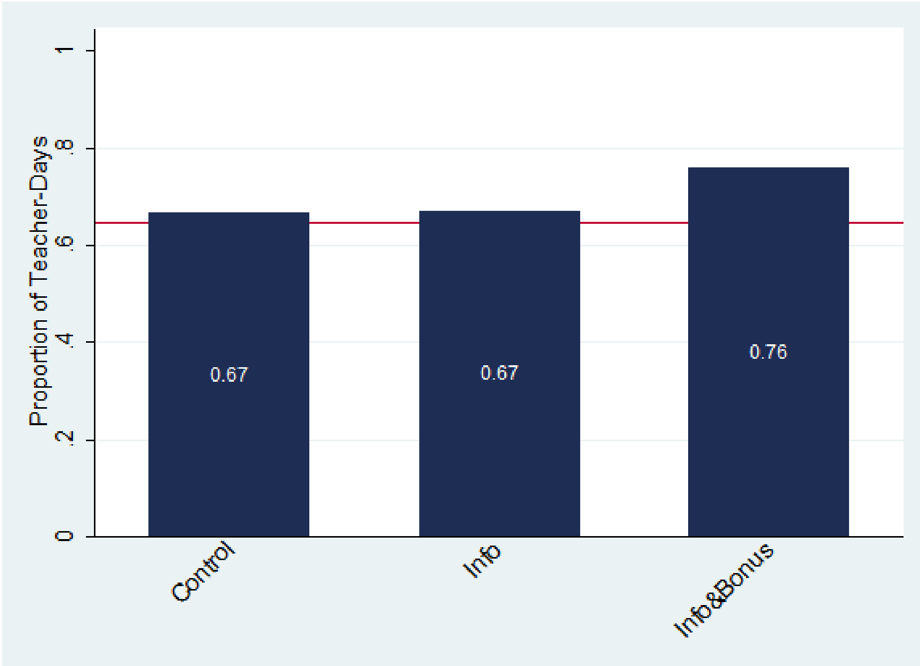
Results are illustrated in Figure B.1, which shows teacher presence levels by treatment arm and by trimester of exposure. Impacts are remarkably stable across the duration of the experiment.

Appendix B.2 Attendance effects for teachers only

Since the theoretical results apply to the attendance of teachers only and not the head teachers, it is important to confirm if the results hold if we restrict analysis of section 4 to only teachers.

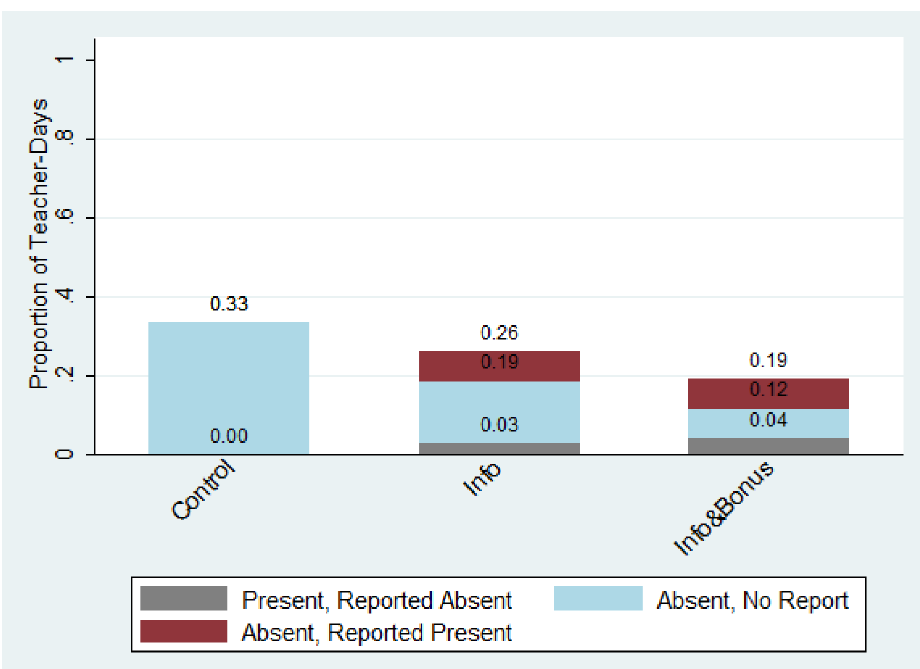
Figures below show that magnitude of the impacts are almost exactly the same. For example, in both cases the probability of making a policy mistake in the *Info&Bonus* schools is 14 and 7 percentage points lower relative to the Control and *Info* schools respectively. Furthermore, all the main results remain statistically significant.

Figure B.2: Teacher Presence



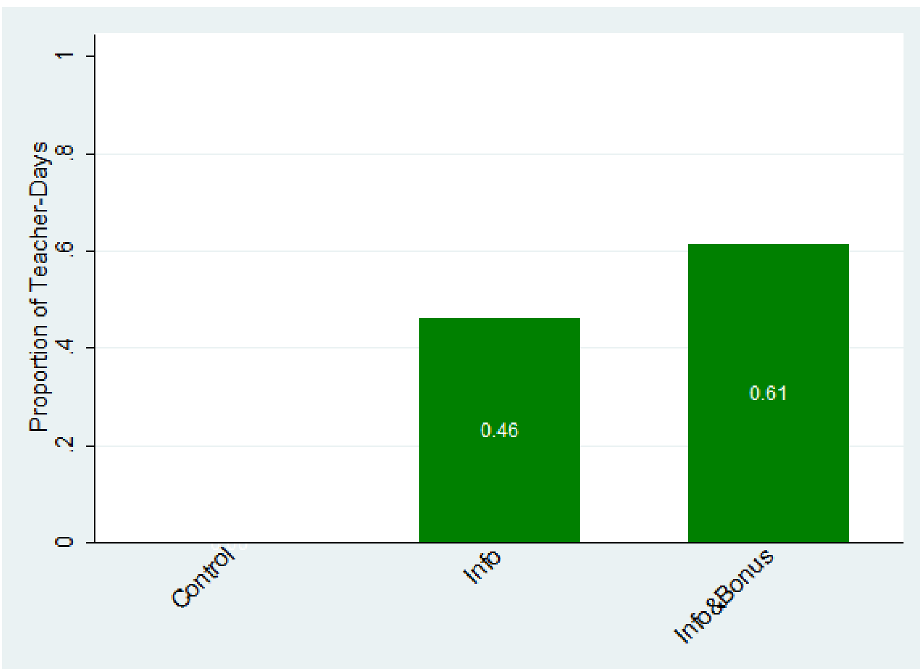
Note: Figure is based on 3,292 teacher-days with independent spot-checks.

Figure B.3: Policy Mistakes



Note: Figure is based on 3,292 teacher-days with independent spot-checks.

Figure B.4: Reported Teacher Attendance



Note: Figure is based on 3,292 teacher-days with independent spot-checks.

Appendix C Detailed Welfare Analysis

This section outlines in more detail the calculations and assumptions underlying the welfare analysis presented in Table 5. As a first step we use two different data sources—administrative data on enrollment and repetition, and a tracked cohort of pupils surveyed before and after the program—to get a conservative estimate on the increase grade attainment in *Info&Bonus* relative to *Info* schools. Next, we use additional data sources to calculate the increase in the net present value of future life-time earnings due to an increase in schooling.

Appendix C.1 Moving from Enrollment to Grade Attainment

To what extent can we attribute the higher enrollment observed in *Info&Bonus* schools to higher grade attainment? Modelling attainment requires two types of assumptions: The (i) persistence of the program and the (ii) persistence of the program's impacts on attainment. For both cases we take the most conservative approach. First, we model the welfare comparison for the actual experiment as it was conducted: a policy intervention to be conducted for one year, with a return to the status quo and an end to project expenditures immediately following. Second, we assume that the expected attainment remains the same for all pupils, except for those who would have dropped out were it not for the program—the averted dropouts. Furthermore, these pupils go on to drop out immediately after withdrawal of the treatment.[23] They therefore only gain one more year of education. These conservative assumptions, again, allow us to estimate the lower bound for welfare analysis.

Given this approach, we now to turn to estimating the averted drop-outs. The two data sources allows for two different strategies, each with different identifying assumptions on transfers. First, using our survey data and combining it with administrative data of repetition figures in 2011 and 2012, we can back out the implied number of drop-outs in grade g and year t, $\Delta_{g,t}$.

The number of pupils, $\pi_{g,t}$, enrolled in grade g at period t, either drop out of schools at the end of the year ($\Delta_{g,t}$), repeat the grade ($\rho_{g,t}$), transition to the following grade ($\tau_{g,t}$), or transfer to another school. Enrollment can therefore be decomposed into the following:

$$\pi_{g,t} \equiv \Delta_{g,t} + \rho_{g,t} + \lambda_{g,t} + \tau_{g,t}$$

Similarly, those enrolled in grade $g+1$ at the beginning of year $t+1$ either repeated the grade, progressed from the previous grade, or transferred from another school ($\varphi_{g,t}$):

$$\pi_{g+1,t+1} \equiv \tau_{g,t} + \rho_{g+1,t} + \varphi_{g,t}$$

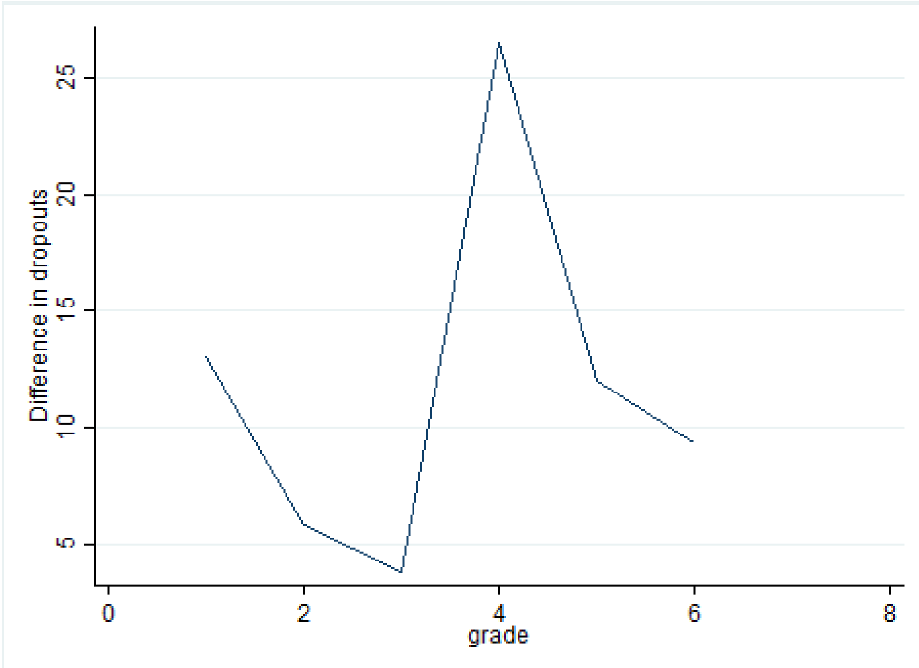
Solving for the above two equations we get the following:

$$\Delta_{g,t} \equiv (\pi_{g,t} - \rho_{g,t}) - (\pi_{g+1,t+1} - \varrho_{g+1,t}) - \mu_{g,t}$$

where we define the net outbound transfers as: $\mu_{g,t} \equiv (\lambda_{g,t} - \varphi_{g,t})$ We have access to administrative

[23] It is plausible that the expected future attainment of these pupils is lower than the average pupil, because they were at the margin of dropping out

Figure C.5: The difference in dropouts per grade between *Info* and *Info&Bonus* schools



Note: Dropouts are calculated using equation 19, using administrative data on enrollment and repetition

data on enrollment and repetition numbers per grade per school, but do not have data on inbound or outbound transfers and can therefore only estimate:

$$\hat{\Delta}_{g,t} = (\pi_{g,t} - \pi_{g+1,t+1}) - (\varrho_{g,t} - \varrho_{g+1,t}) \tag{19}$$

For the difference in $\hat{\Delta}_{g,t}$ across treatment regimes to provide a true estimate of the impact of the program, we need to assume that the net transfers are on average the same across treatment arms. There is a risk of over-estimating the impact on averted drop-outs, for example, if more pupils transfer to the *Info&Bonus* schools because of the program.

Figure C.5 shows the difference in drop-outs between *Info&Bonus* and *Info* for each grade, estimated using equation 19. Note that this difference is highest in grade 4, precisely the grade after which where there is a large drop in enrollment (Figure 5). On average 70 more pupil dropped out in the *Info&Bonus* schools relative to *Info* and this difference is statistically significantly different from zero at the 5% level.[24]

As a second strategy we can derive implied annual dropout and repetition rates from the differences in reduced-form dropout and repetition probabilities observed for the sample tracked from 2010, when the tracked sample of P3 pupils was established, to 2013, when they are observed post-intervention.[25] Data are also available on the grade in which these pupils are enrolled (if any). To back out annual dropout and

[24]Results from the regression analysis, estimated using Equation 15, are available upon request. The number of observations have dropped from 85 to 82, because repetition numbers in 2012 is missing for three schools.

[25]Data on the enrollment status of this sub-sample of pupils are not available for the 2012 baseline to the present study.

Table C.1: Estimated annual pupil transition probabilities, by treatment regime

Treatment, w	Dropout, $\hat{\delta}_w$	Repetition, $\hat{\varrho}_w$
Control	0.299	0.044
Info	0.143	0.060
Info & Bonus	0.0123	0.063

enrollment probabilities, we make note of the fact that three academic years were completed between the time this sample of pupils was drawn and the endline survey, but only one of these years was spent under treatment. Letting δ_w denote the probability of dropout under treatment regime w, and making use of the balance implied by the experimental design, this implies, e.g., that the probability of dropout in the control arm is $\Pr[\text{Dropout}|w = \text{Control}] = 1 - (1 - \delta_{\text{Control}})^3 \approx 1 - 0.344$, using the observed cumulative dropout probability in the Control arm from Table 4. To generate observed probabilities of dropout with one period of treatment exposure in either of the treatment arms $w \in \{\text{Info}, \text{Info\&Bonus}\}$, the implied annualized dropout rates for those arms are given by setting the corresponding observed dropout probability in treatment arm w equal to $\delta_{\text{Control}} + (1 - \delta_{\text{Control}})\delta_{\text{Control}} + (1 - \delta_{\text{Control}})^2\delta_w$. Resulting implied annual transition probabilities are given in Table C.1.

Relative to the estimates of Table 4, dropout rates are lower since these represent annual rather than cumulative dropout probabilities, while differences across treatment arms are exaggerated, since—by virtue of the random assignment of treatment—in expectation all observed differences are attributable to the one year under treatment.

Note that the two different methods discussed above make use of two different identifying assumptions for estimating the true difference in dropouts. The decomposition using administrative data on enrollment and repetition requires that the *net transfers* are not different between the treatment arms. On the other hand, to derive dropout rates using the tracked sample, we need to assume that *inbound transfers* are not affected by the program.

These reduced-form experimental results estimate the short-term enrollment and dropout impacts of assignment to alternative treatment regimes, providing evidence of a statistically and economically significant impact of the *Info & Bonus* arm relative to *Control* and to *Info* alone. With our conservative approach we assume that each averted dropout amounts to no more than an additional year of attainment. But what is the financial impact for those students who remain in school for one more year? We turn to this below.

Appendix C.2 Moving from Attainment to Earnings

As a final step in the welfare analysis, we place a financial value on the increase in grade attainment. As a starting point we want to estimate the net present value of lifetime earnings, NPV_s, of an individual who dropped out after reaching grade s. With the 2011/12 Ugandan National Panel Survey data, we estimate the expected lifetime evolution of wages, $w_{s,t}$, for someone currently residing in one of the size districts where our study took place and who dropped out after completing grade s, using the standard Mincer earnings function (and further restricting the sample to individuals who have obtained at most grade 7 and are 60 years or younger):

$$\ln w_{s,t} = \beta_0 + \beta_1 e + \beta_2 e^2 + \beta_3 s + \varepsilon_{s,t} \tag{20}$$

where $e = t - s$ is the number of years of experience. Since not all of our sample are expected to be in the formal economy, we also want estimate the expected probability, $P_{s,t}$, of earning a wage:

$$P_{s,t} = \alpha_0 + \alpha_1 e + \alpha_2 e^2 + \alpha_3 e^3 + \alpha_4 s + \varepsilon_{s,t} \tag{21}$$

We chose a linear probability model (rather than a logit or probit) for ease of calculating the marginal impact of an additional year of education or experience on the probability of being employed (more on that below). However, even though we use a linear probability model, the probability of formal employment always remains between zero and one.

Table C.2 shows the results of these two regressions.

Table C.2: Enrollment

	(1) Prob. Employed	(2) Log Wage
Years of Schooling	0.00817***	0.0487**
	(0.00)	(0.02)
Years of Experience	0.0114***	0.0538**
	(0.00)	(0.02)
Years of Experience - Squared	-0.000426***	-0.000826*
	(0.00)	(0.00)
Years of Experience - Cubed	0.00000452***	
	(0.00)	
Constant	-0.0326***	13.26***
	(0.01)	(0.25)
Observations	5752	282

Note: Columns (1) and (2) reflect regression outputs from equations 20 and 21 respectively. In both regressions the sample is restricted to individuals who have obtained at most grade 7, are no longer in school, and are younger than 60 years. In column (2) the sample is further restricted to individuals who earn a salary. Standard errors are in parentheses. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

Note that equations 20 and 21 need not make any causal claims on the returns to education. At this point we do not wish to estimate the impact of schooling on earnings; we merely want to estimate the right counter-factual—the expected net present value of future life-time earnings for each level of primary school grade attainment.

Using the predicted evolution of wages and employment probability over time, $\hat{w}_{s,t}$ and $\hat{P}_{s,t}$, and assuming that everyone stops working at age 60, we can now estimate the lifetime earnings of individual who dropped out at the end of grade s:

$$N\hat{P}V_s = \sum_{t=0}^{t=53-s} \left(\frac{1}{1+r}\right)^t \left(\hat{P}_{s,t} \cdot \hat{w}_{s,t} + \left(1 - \hat{P}_{s,t}\right) \cdot A\right) \tag{22}$$

To simplify analysis, we assume that there are only two sectors, formal wage employment and subsistence agriculture. We further conservatively assume that earnings from agriculture, A, does not depend on years of education and experience.[26] To estimate A, we take the average agricultural income from the sample of rural households whose main source of income is agriculture and divide this by two.[27]

Next, we calculate the *change* in the NPV_s due to one more year of education. For this, we need an estimate for the causal impact of education wage and the probability of employment. We use 6.8% for a lower bound of the causal impact of education on earnings. This is the most conservative estimate from studies that credibly estimated the causal impact of education on earnings in developing countries (Card 2001, Duflo 2001).[28] As an upper bound we use 15.9%, the value calculated by Montenegro and Patrinos (2014) for Uganda.

To the best of our knowledge, no studies in developing countries have looked at the impact of education on gaining formal employment. As a best guess, we rely on our own estimates in equation (2), $\hat{\alpha}_1 = 0.008$. In table 5 in the welfare section we also show results under the most conservative assumption of $\hat{\alpha}_1 = 0$.

Using these numbers, we can now calculate the welfare gain of one additional year of education for someone who would have dropped out after s years of school:

$$\frac{\partial NPV_s}{\partial s} = \sum_{t=0}^{t=53-s} \left(\frac{1}{1+r}\right)^t \left(0.068 \cdot \hat{w}_{s,t} \cdot \hat{P}_{s,t} + 0.008 \cdot (\hat{w}_{s,t} - A)\right) \tag{23}$$

Table C.3 brings together our estimates of averted dropouts and the returns to education to calculate a total (per school) financial gain due to the introduction of the bonus scheme (comparing *Info&Bonus* with *Info* schools). Column (2) shows the result for equation 22 (setting $r = 10$): the expected net present value of future expected earnings, for each level of grade attainment. Column (3) shows the results

[26] It is plausible that productivity in the non-formal sector also improves with education, but to our knowledge there are no studies in developing country context that credibly estimate this. Furthermore, in our sample, experience and education is not significantly correlated with household earnings from subsistence farming.

[27] This is plausibly an over-estimate of the individual earnings from agriculture (and thus leads to a more conservative estimate of the impact of the program), because more than two people per household typically work on the household's farm. Another approach would be to divide by the number of household members who claim to have worked on the household farm in the past year (roughly 4).

[28] Duflo (2001), using a difference-in-difference approach based on year of enrollment and distance from newly constructed schools, estimated a return to education ranging between 6.8 and 10.6 per cent in Indonesia. Maluccio (1997) estimated 11.3% in Philippines. Other well-identified studies that use samples from developing countries (Card 2001) typically find larger returns to education.

from equation 23: the expected improvements in NPV due to one additional year of schooling complete (assuming a causal impact of 6.8% and 0.8% on wage earnings and probability of formal employment respectively). Columns (4) and (6) show the results from section C.1: the average reduction in drop-outs per grade per school, due to the introduction of financial incentives. In column (3), this is calculated using the administrative data on enrollment and repetition; in column (5) this is estimated using the cohort data. Finally, columns (4) and (6) show the average financial gain per grade per school.

Table C.3: Difference in average per school financial returns between *Info&Bonus* and *Info* schools

			Administrative Data		Cohort Data	
Grade	NPV	Gain	Averted Dropouts	Total gain	Averted Dropouts	Total gain
1	2730.71	30.93	14.02	433.62	12.56	388.52
2	2756.32	35.85	7.091	254.23	9.76	349.88
3	2785.85	41.13	5.857	240.89	10.62	436.72
4	2819.53	46.78	27.64	1292.89	10.80	505.12
5	2857.64	52.82	12.31	650.19	10.44	551.19
6	2900.43	59.28	7.117	421.87	8.27	490.50
Total				**3293.69**		**2721.94**

Note: Column (1) shows the expected net present value of future expected earnings, given each grade attainment, assuming a discount rate of 10%. Column (2) shows the gain in NPV due to an additional year of schooling achieved, assuming a causal impact of 6.8% and 0.8% respectively on wage earnings and probability of formal employment. Columns (3) and (5) indicates treatment effect - the average number of averted dropouts per grade per school due to the program - calculated using the two different data sets. Columns (4) and (6) show the average financial gain per grade per school.