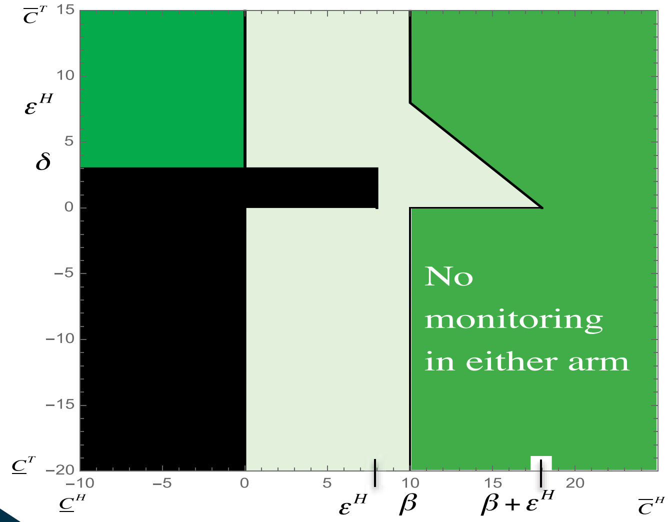
Towards Sustainable Development

OCCASIONAL PAPER. 40



PAY FOR LOCALLY MONITORED PERFORMANCE?

A Welfare Analysis For Teacher Attendance In Ugandan Primary Schools

August 2016

Jacobus Cilliers, Ibrahim Kasirye, Clare Leaver, Pieter Serneels, and Andrew Zeitlink

Occasional Paper No.40

PAY FOR LOCALLY MONITORED PERFORMANCE?

A Welfare Analysis For Teacher Attendance In Ugandan Primary Schools

August 2016

The Economic Policy Research Centre (EPRC) is an autonomous not-for-profit organization established in 1993 with a mission to foster sustainable growth and development in Uganda through advancement of research –based knowledge and policy analysis. Since its inception, the EPRC has made significant contributions to national and regional policy formulation and implementation in the Republic of Uganda and throughout East Africa. The Centre has also contributed to national and international development processes through intellectual policy discourse and capacity strengthening for policy analysis, design and management. The EPRC envisions itself as a Centre of excellence that is capable of maintaining a competitive edge in providing national leadership in intellectual economic policy discourse, through timely research-based contribution to policy processes.

Any enquiries can be addressed in writing to the Executive Director on the following address:

Economic Policy Research Centre
Plot 51, Pool Road, Makerere University Campus
P.O. Box 7841, Kampala, Uganda
Tel: +256-414-541023/4
Fax: +256-414-541022
Email: eprc@eprc.or.ug
Web: www.eprc.or.ug

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[‖]

August 2016

Abstract

Public sector organizations often rely on reports by local monitors that are costly to verify and that serve twin objectives: to incentivize agent performance, and to provide information for planning purposes. Received wisdom has it that pay for locally monitored performance (P4LMP) will result in collusion and undermine both objectives. But simple Coasian logic suggests the reverse: P4LMP puts transferable money on the table and may enable interested parties to bargain to a more efficient outcome. This paper develops a theoretical model that shows why, and for which parameters, the welfare-enhancing Coasian scenario exists. Focusing on education, we model how the preferences of a teacher (agent) and head-teacher (local monitor) affect actual and reported teacher attendance, and how these equilibrium outcomes depend on the financial stakes attached to reports. To capture the value of information, we also consider the welfare of a bureaucracy that makes a costly policy mistake when holding the wrong belief about teacher performance. We test the model and estimate the predicted effects using data from a field experiment in Ugandan primary schools, randomly varying whether head teachers' reports of teacher attendance are tied to bonus payments or not. Consistent with Coasian logic, P4LMP increased actual and reported teacher attendance (by 9 and 15 percentage points respectively) and reduced policy mistakes (by 7 percentage points) relative to unincentivized local monitoring. We use these experimental impacts to undertake a detailed cost-benefit analysis and conclude, even under conservative assumptions, that welfare improved when paying for locally monitored performance.

JEL Classification Numbers: D61 H52 I25 I26

---

*We are grateful for useful feedback from David Evans, James Habyarimana, Julien Labonne, and Sandip Sukhtankar. We thank seminar participants at the Tinbergen Institute, Universitat Autònoma de Barcelona, World Bank, University of Oxford, Georgetown University, University of Bath, the Center for Global Development, University of East Anglia, Wageningen University and the following conferences: AEA Annual Meeting, RES Annual Conference, IGC Growth Week, World Literacy Summit, Uganda Education Research Symposium, SEEDC, IiG, NEUDC, CMPO and Non-Profits, Governments, and Organizations. The 'Improving Institutions for Pro-Poor Growth' Research Consortium, the IGC, and the Oxford Institute for Global Economic Development provided funding. All errors and omissions are our own.

[†]Georgetown University

[‡]Economic Policy Research Centre

[§]Blavatnik School of Government, University of Oxford and Centre for Economic Policy Research

[¶]University of East Anglia and IZA

[‖]Georgetown University and Center for Global Development

1 Introduction

Public sector organizations around the world rely on reports by local monitors that are costly to verify. Typically, these reports serve two objectives: to incentivize desired behavior, and to provide information for planning purposes. To these ends, in many education systems head-teachers submit pupil enrollment and attendance figures, and schools (sometimes even pupils) receive financial transfers based on these reports. In health systems, it is common for hospital administrators to submit performance indicators such as the number of patient visits or hospital waiting times, and for health-care professionals to be rewarded based on these reports. Governments use such reports not only to incentivize agents but also to make policy decisions in aggregate, for example relating to facility construction, human resource transfers, the taxation of unhealthy habits, and public health campaigns.

When stakes are attached to these reports—whether pecuniary or reputational—there is a clear risk of mis-reporting. Across 21 countries in Africa, head-teacher over-reporting of pupil enrollment figures increased dramatically when countries introduced school funding on a per-pupil basis (Sandefur & Glassman 2015). Veterans Affairs hospitals in the US kept patients off official waiting lists in order to meet targeted 14-day waiting times for appointments (VA Office of Inspector General 2014). These distortions not only weaken incentives for providers, but also undermine governments' ability to plan and allocate resources effectively.

Problems can also arise when an agent, rather than local monitor, stands to gain from mis-reporting. In Kenya, head-teachers were asked to monitor teacher presence and reward teachers based on these reports. Head-teachers systematically overstated teacher presence and there was no improvement in teacher performance (Chen, Glewwe, Kremer & Moulin 2001). Similarly, in India, teachers could reward their pupils for attending school and were found to manipulate student presence figures (Linden & Shastry 2012). Environmental auditors, when hired by the firms they investigated, systematically understated the extent of pollution (Duflo, Greenstone, Pande & Ryan 2013). These examples are suggestive of collusion: the local monitor lies about agent performance in return for a share of the reward.

Is collusion between local monitors and the targets of bureaucratic incentive schemes inevitable? Conventional wisdom suggests as much. Campbell's Law, which holds that "the more any quantitative social indicator is used for decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor", has guided much thinking on accountability in schools and other domains of public sector organization (Campbell 1979, Rothstein 2011, Neal 2013).

However, familiar Coasian logic suggests that Campbell's law need not always hold. Absent transaction costs, parties interested in service delivery outcomes (parents, head-teacher, teaching staff and government officials, say, in an education context) can bargain to an efficient allocation of delivery effort. Side payments allow frontline agents to internalize the social benefit of service provision, alongside their private cost of effort. As others have observed (e.g. Dixit 1996), if service delivery outcomes are inefficiently low, then transaction costs must be preventing the interested parties from bargaining effectively. Such frictions are widespread in low-income settings: physical distances and/or a lack of comprehension may impede information flows between parties, while financial constraints may limit the scope for transferable utility (Banerjee, Banerji, Duflo, Glennerster & Khemani 2010). And observable measures of effort, such as presence, are correspondingly low (Chaudhury, Hammer, Kremer, Muralidharan & Rogers 2006). It follows that a policy that reduces these frictions, for instance by making payments based on local monitoring and thus putting transferable resources on the table, could improve the efficiency of service delivery, precisely because of (not despite) the role played by side payments.

This paper sets out to answer three related questions at the heart of payment for locally monitored performance—hereafter referred to as P4LMP—in the context of public service delivery. Can P4LMP induce improvements in service providers' behavior? Does P4LMP reduce or improve the quality of reported information for planning purposes? And what is the overall impact of P4LMP on social welfare?

We address these questions in the context of Ugandan primary education, where we explore how incentives for primary school teachers can be effectively designed when based on local monitoring by head-teachers. This is an important issue in Ugandan education: 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).[1] Remote school locations and limited resources for inspections make local monitors a particularly important source of information on school inputs in this context.

Our analysis combines a theoretical model with empirical results from a field experiment and a representative household survey. To begin, we model how the preferences of both teacher (agent) and head-teacher (monitor) affect teacher attendance and head-teacher monitoring and reporting, and how these equilibrium outcomes depend on the financial stakes attached to the reports. To evaluate the potential trade-off between performance and quality of information, the model also considers the welfare of a bureaucracy that values teacher presence but makes a costly policy mistake if it holds the wrong belief about teacher absence.[2] Next, we conduct a field experiment to provide estimates of the key behavioral effects identified in the theory. Finally, using a welfare criterion suggested by the theory, we combine our experimental estimates with additional data to undertake a cost-benefit analysis of the alternative intervention designs.

The theoretical model delivers three predictions. Relative to unincentivized local monitoring, P4LMP increases the probability of actual teacher attendance, increases the probability of reported teacher attendance (and hence a bonus payment), and has two competing effects on the probability of a bureaucratic policy mistake which leave the overall impact ambiguous. These predictions are driven by both the collusive story and the Coasian logic alluded to above. P4LMP introduces a source of transferable utility between head-teachers and teachers, who use this to bargain to locally efficient outcomes.

The intuition runs as follows. P4LMP puts money on the table. The head-teacher, who plausibly has the bargaining power, can either use this transferable utility to induce the teacher to attend, or collude and file a false report. If the head-teacher cares about teacher attendance, even slightly so, then when the costs of attendance and monitoring are not too high, it will be mutually beneficial for the head-teacher and teacher to agree on a side-contract where the head-teacher effectively 'pays' the teacher to attend. However, when the cost of attendance is high but the cost of monitoring is not, the head-teacher opts to collude and file a false report. Hence there are more truthful reports (Coasian logic) but also more false reports of teacher presence (collusion). These two forces also explain why the effect on the quality of

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

[2] Since on average teachers are present more than 50 percent of the time, the bureaucracy's "best bet" under no information is that the teacher is present. Hence rational, Bayesian bureaucratic behavior results in a policy mistake when teacher absence is unreported.

bureaucratic decision-making has an ambiguous sign. On the one hand, there are parameter values where the bureaucracy receives a truthful report of teacher presence rather than no report of teacher absence, thereby averting a policy mistake. On the other hand, there are also parameter values where it receives a false report of teacher presence instead of a truthful report of teacher absence, thereby inducing a policy mistake.

The model suggests a natural welfare criterion, namely the bureaucracy's payoff, which is higher when the teacher attends and lower when a bonus is paid out and/or a policy mistake is made. Consistent with received wisdom, there are parameter regions where P4LMP reduces the bureaucracy's (ex post) payoff. However, there are also parameter regions where P4LMP *increases* this payoff, namely when the cost of attendance is intermediate and the cost of monitoring is not too high. Here P4LMP incentivizes the desired behavior and can also provide unbiased information for policy purposes. The key contribution of the theory is to show why and where this welfare-enhancing scenario exists.

Next, in order to test the theoretical predictions of the model and also deliver a full (ex ante) welfare comparison, we conducted a randomized controlled trial, experimentally varying the existence of financial stakes attached to local reports. This allows us to estimate the sign and magnitude of the three behavioral effects. In one treatment (20 schools), head-teachers were requested to submit reports of teacher attendance using mobile technology. This information was then collated and relayed back to the community. The second treatment, our P4LMP arm (25 schools) was exactly the same, except that teachers received bonus payments (roughly 30 percent of their monthly salary) if they were reported as present regularly over a month. Another forty schools were randomly assigned to a control.

We conducted our own independent spot checks of teacher presence—both prior to the intervention and during every term that the intervention took place—which we then compared to head-teacher reports. A school survey captured basic school and teacher characteristics. We also measured learning outcomes and grade retention for a cohort of students that we tested before and after the intervention.

Our empirical results are broadly in line with the predictions of the model. With the introduction of bonus payments, both teacher attendance and reported teacher attendance are higher and policy mistakes due to non-reports of teacher absence are lower. There is also evidence of collusion—absent teachers falsely reported as present—in both treatment arms. However, relative to unincentivized local monitoring, P4LMP *reduces* policy mistakes.

In addition, we find a substantial increase in pupil enrollment with the introduction of P4LMP. This trend is observed across all grades, but is highest in grades where school dropouts are a serious problem. The large changes in pupil composition preclude any conclusive claim of impacts on learning.[3]

We use these results to undertake a cost-benefit analysis of the alternative intervention designs. In order to place a financial value on the expected total pupil benefit from improved teacher performance, we back out the implied dropout rate in the different treatment arms, using both administrative data on enrollment and repetition, and our own data on retention from the cohort of students surveyed prior to the intervention. We then calculate the financial returns to an additional year of education in Uganda, using earnings and employment data from the Uganda National Panel Survey and previous authors' estimates of the rate of return to education in developing countries. Our school-level estimates of the expected total pupil benefit from P4LMP range from USD $1,745$ to USD $6,122$, depending on assumptions and

[3]Performing Lee bounds, the impact is either positive or negative, depending on the assumptions for sample selection.

data sources. Our preferred estimate, based on conservative assumptions, is USD $2,722$. Since this figure exceeds the school-level bonus cost of USD $2,250$, it follows that P4LPM passes a cost-benefit test even before considering the quality of information (recall the likelihood of policy mistakes actually fell). Under our most conservative assumptions, P4LMP passes a cost-benefit test only if the welfare gain from avoiding a policy mistake is sufficiently large—specifically, USD 40 per school per day. To provide a benchmark, the Government of Uganda has set a target to spend no more than USD 44 per school inspection (Ministry of Education, Science, Technology and Sports 2014). If we take this as the Government's valuation of information, then we can conclude that even under our most conservative assumptions P4LMP was welfare-enhancing in the context of Ugandan primary education.

The key takeaways from our analysis are that, contrary to received wisdom but in line with simple Coasian logic, P4LMP can improve both service delivery and the quality of information for planning purposes. Moreover, this dual objective can be met sufficiently cheaply to pass a cost-benefit test.

This paper contributes to three aspects of the literature on state effectiveness in poor countries. First, a number of papers have sought to understand how incentives—both pecuniary and non-pecuniary—impact the effort levels of frontline service providers. Researchers have typically collected the performance metric themselves, whether adminstering student assessments to measure outcomes of provider effort (Muralidharan & Sundararaman 2011), or administering tamper-proof disposable cameras to measure teacher presence (Duflo, Hanna & Ryan 2012). Such experiments provide proof of concept, demonstrating a necessary condition for impacts: that agents respond to the performance incentive when ideally administered. But recent attempts to extend 'automated' monitoring to the public sector have proven challenging, notably in health in India (Banerjee, Duflo & Glennerster 2007, Dhaliwal & Hanna 2014) and in education in Haiti (Adelman, Blimpo, Evans, Simbou & Yarrow 2015). These experiences underscore the importance of studying, as we do, how monitoring contracts and technologies interact with the preferences of local parties.[4]

Second, a small but growing literature documents the prospects of digital technologies to improve public service delivery.[5] As with mobile money, such technologies offer opportunities to circumvent frictions that otherwise lead to market failures (Suri, Jack & Stoker 2012). Callen and coauthors (2016) demonstrate that information collected by smartphones (in place of paper forms) on health worker absence 'crowds in' central inspections in politically competitive constituencies. Aker and Ksoll (2015) find that phone calls by government officials to local parties (teachers, community representatives, and a random sub-sample of students) improve learning outcomes of an adult education program in Niger. Our paper contributes to this, hitherto empirical, literature by studying a technological intervention that is (a) motivated by theory and (b) utilizes local monitoring coupled with financial incentives.

Third, there is growing interest in applying insights from the public finance literature, particularly welfare analysis, to questions in development economics. Baird, Hicks, Kremer & Miguel (2015), for example, calculate the long-term financial gain due to improved health of children that received de-

[4]This is also a key theme of recent work by Duflo and colleagues (2013), who show that environmental auditors suffer from conflicts of interest that affect the data they report and the subsequent behavioral responses of firms.

[5]As the 2016 World Developement Report notes, digital technologies can help improve service delivery by: informing citizens; streamling processes; receiving feedback; and "improving service provider management through better monitoring so that government workers both show up at work and are productive" (WDR 2016, p. 157). Our paper is part of the literature documenting the fourth of these so called 'digital dividends'.

worming in Kenya. They argue that the additional tax revenue from future income alone is sufficient to pay for the program. Similar approaches have been taken in recent work on tax policy (Best, Brockmeyer, Kleven, Spinnewijn & Waseem 2015) and unemployment benefits (Gerard & Gonzaga 2014). Our paper speaks to this interest, deploying an explicit theoretical framework to guide empirical analysis and welfare comparisons.

The remainder of the paper proceeds as follows. Section 2 introduces a theory of local monitoring that yields both positive predictions for monitors' and teachers' responses to alternative treatments and normative, welfare criteria for comparing outcomes under each. Section 3 outlines the field experiment and data. Section 4 uses this data to estimate the sign and magnitude of the behavioral effects predicted by the theory. 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 offer concluding remarks, including scaling up the intervention in Uganda and external validity.

2 Theory

The model provides a stylised way to analyse the effect of local monitoring and incentives. We consider three variants which, foreshadowing sections below, we refer to as experimental arms. The basic structure in all arms is that teachers can choose between showing up for work or not. In the *Control* arm, teacher attendance remains unmonitored. Having pinned down a baseline, we then study how the introduction of local monitoring by the head-teacher, who reports to a government bureaucracy, affects teacher attendance. We refer to this as the *Info* arm. Next, we investigate how combining local monitoring with financial incentives triggered by the head-teacher's reporting affects teacher attendance, referring to this as the *Info & Bonus* arm. We compare equilibrium outcomes—teacher attendance, head-teacher monitoring/reporting (and hence bonus payments), and bureaucratic decision-making—as well as the bureaucracy's ex post and ex ante payoffs across the three experimental arms.

2.1 Model

Players and actions The economy consists of a teacher (he), a head-teacher (she), a government bureaucracy (it), and n identical pupils. 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 sends a *truthful report* iff $r = a$. If the head-teacher does not monitor, $m = 0$, she cannot send a report.[6] 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, $r = 1$. Pupils play no active role under any treatment arm.

[6]We provide evidence to support this assumption that monitoring is necessary to send a report in Section 3 below.

Payoffs All players are risk neutral. Net of any side-transfers, the payoffs to 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 pupil 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 pupil benefit from teacher attendance less the cost of the cash bonus and/or a 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 $-\underline{C}^T > \overline{C}^T > 0$ (so that attendance in Control schools is more than 50%). 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.[7] 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

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

implying that the pupil gain from teacher attendance must be sufficiently high. As we show in Appendix A, this is a sufficient condition for the bureaucracy to choose a cash bonus satisfying $\beta \geq \varepsilon^H$ in the Info & Bonus treatment. Focusing on this case simplifies the exposition of the results.

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:

[7] 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.

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.

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 is:

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 Analysis

We present our theoretical results graphically, relegating the formal details to Appendix A. Referring to Figures 1 to 3, we state equilibrium outcomes for teacher attendance, head-teacher monitoring/reporting (and hence bonus payments), and bureaucratic decision-making. Collecting these outcomes together, we then compare the bureaucracy's *ex post* payoff (realized after attendance and monitoring costs have been drawn by nature and players have made their moves) across experimental arms. Since these cost realizations are unknown to all but local players, our welfare criterion is the bureaucracy's *ex ante* payoff (taking expectations over costs and anticipating equilibrium play). Our analysis concludes with a welfare ranking of the experimental arms, based on the bureaucracy's ex ante payoffs.

Teacher attendance Figure 1 plots realisations of the monitoring cost, C^H, and attendance cost, C^T, for which the teacher chooses to attend school. Here, as in subsequent figures, the remaining parameters are held fixed at $\varepsilon^H = 8, \delta = 3$, and $\beta = 10$. Naturally, the teacher chooses to attend if he derives an intrinsic benefit from doing so. The area below the $x-$axis therefore represents the baseline probability of attendance in the Control arm. The dark blue shading shows the region of the parameter space where the teacher attends in the Info arm but is absent in the Control arm; i.e. where the local monitoring technology changes the teacher's behaviour from absence to attendance. The light blue shaded area is where the teacher attends in the Info & Bonus arm but is absent in the Info and Control arms. In this parameter region, the availability of financial incentives in addition to the local monitoring technology changes the teacher's behaviour from absence to attendance. We refer to this as the *attendance effect* of local monitoring and incentives.

The intuition behind Figure 1 is straightforward. Teacher attendance is higher in the Info arm than in the Control arm because the head-teacher is able to leverage the reputation cost δ to secure the benefit ε^H. If δ exceeds the attendance cost C^T, and ε^H outweighs the monitoring cost C^H, then the players reach an equilibrium where the head-teacher commits to monitor and report truthfully, and the teacher attends. Teacher attendance is higher still in the Info & Bonus arm because there is now transferable utility on the table, in the form of the incentive bonus β. The head-teacher can either use this transferable utility to induce the teacher to attend, or she can collude and file a false report. If the attendance cost is intermediate—higher than δ but less than ε^H—*and* β and ε^H together exceed the combined attendance and monitoring costs (a joint efficiency requirement), the head-teacher opts for the former outcome. The players reach an equilibrium where they first agree on a side-contract that commits the head-teacher to monitor and report truthfully and the teacher to share some fraction of the bonus, and the teacher subsequently attends. Hence, the attendance effect stems from the fact that the bonus β is both bigger and more transferable than the reputation cost δ.

Head-teacher monitoring/reporting and bonus payments Figure 2 plots cost realizations where the head-teacher chooses to monitor and report the teacher absent, or present (thereby triggering the bonus). Clearly, the head-teacher chooses to monitor if she derives an intrinsic benefit from doing so. But she may also choose to monitor to induce teacher presence. In the Info arm, this happens in the area above the $x-$axis and to the right of the $y-$axis where the local monitoring technology changes the teacher's behaviour from absence in the Control to attendance in the Info arm (i.e. the dark blue region in Figure 1). Lacking any incentive to do otherwise, if the head-teacher chooses to monitor, then she reports truthfully. The black shading shows the region where the head-teacher monitors and reports the teacher present, while the dark green shading shows where she monitors and reports the teacher absent.

In the Info & Bonus arm, the head teacher again chooses to monitor if she derives an intrinsic benefit from doing so and may also choose to monitor in order to induce teacher presence. The latter happens in the area above the $x-$axis and to the right of the $y-$axis where the availability of financial incentives in addition to local monitoring changes the teacher's behaviour from absence in the Control to attendance in the Info & Bonus arm (i.e. the dark blue plus light blue region in Figure 1). The crucial difference now is that the head-teacher has an additional reason to monitor, namely to trigger payment of the bonus. She therefore monitors and reports the teacher present whenever her monitoring cost is less than the bonus, even if this has no causal impact on teacher attendance. It follows that the black, dark green and light green shading together show the region where the head-teacher monitors and reports the teacher present. In contrast to the Info arm, she never monitors and reports the teacher absent.

For what follows, it will be helpful to refer to two effects. The first is the the *reported teacher attendance effect* of local monitoring and incentives, depicted by the sum of the dark and light green regions in Figure 2. Reported teacher attendance is higher in the Info & Bonus arm because there are more truthful reports (via the attendance effect) but also more false reports of teacher presence. The second is the *payout effect* of local monitoring and incentives, depicted by the sum of all shaded regions. Intuitively, the head-teacher can either use the transferable utility β to induce the teacher to attend, or she can collude and submit a false report. If the attendance cost is higher than ε^H, and providing her monitoring cost can be covered by β, the head-teacher opts for the collusive outcome. Hence, the magnitude of both effects is driven by both the bonus β and the direct benefit ε^H.

Bureaucratic decision-making Turning to bureaucratic decision-making, note that there are two separate cases when the (rational and Bayesian) bureaucracy makes a policy mistake. The first case is in the Info & Bonus arm when the head-teacher falsely reports the teacher present. Even though the bureaucracy anticipates that some reports will be false, it is still rational for it to take reported presence at face value and choose $p = 1$ to minimise the *expected* loss from making a mistake.[8] In the event that the teacher is actually absent, the bureaucracy therefore makes a mistake. The second case is when the head-teacher does not monitor and the teacher is absent. In all arms, the bureaucracy's rational response to no report is to presume that the teacher is present.[9]

These two cases underlie the policy mistakes shown in Figure 3. The red shading depicts the region of the parameter space where there is a policy mistake in the Control and Info & Bonus arms but not in the Info arm. In the Control arm, this is because there is no report and the teacher is absent, while in the Info & Bonus arm this is because the teacher is falsely reported present. In contrast, in the Info arm the teacher is truthfully reported absent, enabling the bureaucracy to take the correct decision. The dark green shading depicts the region of the parameter space where the bureaucracy makes a policy mistake in the Control arm (no report and the teacher is absent) but not in either of the treatment arms (due to truthful reporting). Finally, the light blue shading depicts the region of the parameter space where there is a policy mistake in both the Control and Info arms (no report and the teacher is absent) but not in the Info & Bonus arm (truthful reporting). Hence, in this region, the availability of financial incentives actually *improves* the quality of locally collected information.

It follows that the bureaucracy is less likely to make a policy mistake in the two treatment arms than in the Control. To establish whether a policy mistake is more or less likely when financial incentives are available in addition to local monitoring, we must compare the light blue and red areas. We refer to this difference—red minus light blue—as the *decision effect* of local monitoring and incentives. Under the parameter assumptions in Figure 3 ($\varepsilon^H = 8, \delta = 3$, and $\beta = 10$), the light blue area is larger than the red

[8] The posterior probability that the teacher is present conditional on being reported present is greater than 0.5, implying that $p = 1$ is the optimal decision. See Appendix A.

[9] In the Control arm, given $-\underline{C}^T > \overline{C}$, the bureaucracy anticipates that the probability of teacher attendance is greater than 0.5 and so its optimal policy decision is $p = 1$. In the Info and Info & Bonus arms, the bureaucracy anticipates that the probability of teacher attendance will be even higher and so $p = 1$ remains the optimal decision.

area, indicating that Info & Bonus has the informational advantage.

The fact that attaching financial incentives to local monitoring reports could improve the quality of information (negative decision effect) may at first seem surprising. But the intuition is simple and follows from the attendance effect. In the light blue region, the bonus serves as additional transferable utility that enables the head-teacher to both cover her monitoring costs and induce the teacher to attend. Without this transferable utility, there is no report and the teacher is absent, leading to a policy mistake. The flip-side in the red region is that, when it is too costly to 'pay' the teacher to attend but the head-teacher is nonetheless intrinsically motivated to monitor, the bonus gives her an incentive to submit a false rather than truthful report. This lack of truthtelling results in a policy mistake.

Summing up and the bureaucracy's ex post payoff Collecting together the above equilibrium outcomes, we can now compare the bureaucracy's ex post payoff (realized after attendance and monitoring costs have been drawn by nature and players have made their moves) across experimental arms.

Figures 1 to 3 show that there are no cost realisations where the bureaucracy's ex post payoff is lower in the Info arm than in the Control. Either this ex post payoff is strictly higher in the Info arm (the dark blue region of Figure 1 where the gain is $n\varepsilon^P$ from higher teacher attendance and κ from fewer policy mistakes, and the dark green region of Figure 2 where the gain is κ from fewer policy mistakes) or the two arms yield the same equilibrium outcomes and so an identical ex post payoff.

The comparison between Info and Info & Bonus is more complex. Table 1 summarises the differences in equilibrium outcomes across these two treatment arms and highlights six distinct parameter regions (shown in Figure 4). The first point to note is that there are cost realisations where the bureaucracy's ex post payoff is lower in the Info & Bonus arm than in the Info arm. In Region 1, the loss is $\beta + \kappa$ because the introduction of financial incentives has no causal impact on teacher attendance (the teacher is absent in both arms) but results in a bonus payout and poorer bureaucratic decision-making due to false reporting of teacher presence. This corresponds to the received wisdom noted in the Introduction. In Regions 4 and 5, the loss is β because there is no causal impact on teacher attendance (the teacher is present in both arms) or bureaucratic decision-making, but again a bonus is paid out.[10] Crucially, there are also cost realisations where the bureaucracy's ex post payoff is *higher* in the Info & Bonus arm than in the Info arm. In Region 2, the gain is $n\varepsilon^P - \beta$ because the introduction of financial incentives has a causal impact on teacher attendance, although at the expense of a bonus payout. In Region 3, the gain is $n\varepsilon^P - \beta + \kappa$ because additionally there is better bureaucratic decision-making due to truthful reporting of teacher presence. These regions corresponds to the alternative Coasian scenario noted in the Introduction. The contribution of our theoretical model is to show why and where this welfare-enhancing scenario exists.

Expected welfare Recall that our welfare criterion is the bureaucracy's ex ante payoff, taking expectations over C^T and C^H. Since the bureaucracy's ex post payoff is weakly higher in the Info arm relative to the Control for all cost realizations, local monitoring of performance must increase expected welfare relative to the status quo. On the other hand, attaching financial incentives to local monitoring

[10] In Region 4, there is no policy mistake in either arm because the head-teacher truthfully reports the teacher present; in Region 5, there is a policy mistake in both arms because the head-teacher falsely reports the teacher present in the Info & Bonus arm and sends no report in the Control arm.

reports, i.e. moving from Info to Info & Bonus, has an ambiguous effect on expected welfare driven by the contrasting regions highlighted in Table 1. Specifically, *paying* for locally monitored performance improves expected welfare only if:

$$\underbrace{n\,\varepsilon^P \cdot (\text{attendance effect})}_{\text{expected total pupil benefit}} - \underbrace{\kappa \cdot (\overbrace{\text{decision effect}}^{\text{sign?}})}_{\text{expected decision cost}} > \underbrace{\beta \cdot (\text{payout effect})}_{\text{expected bonus cost}}. \tag{2}$$

It follows that, while our theoretical model structures the analysis, it only gives a partial expected welfare ranking. For this reason, in Sections 3 and 4 we turn to the data collected via our field experiment to estimate the sign and magnitude of the three behavioral effects in equation (2). In Section 5, we supplement this with data from the Uganda National Panel Survey to estimate the expected total pupil benefit. Finally, using this estimate, we establish whether paying for locally monitored performance improved expected welfare under reasonable assumptions on the social loss from policy mistakes, κ.

3 Field Experiment

3.1 Intervention and Experimental Design

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

Working with World Vision, we trained head-teachers, assisted by their deputy, in the use of a platform which allowed them to report teacher attendance on a mobile device. This information, combined with a unique identification number of the school, teacher and monitor, was sent to a central database in Makerere University where project staff processed the data. The platform was added to a phone that we provided, and which was kept at the school.[11] 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 (USD 23, or roughly 30 percent of their monthly salary) if they were reported as present every week that month. Adopting the language of Section 2, 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.[12] Stratifying by district, we randomly assigned 40 schools to the control, 20 schools to the Info arm and 25 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.

[11] This is consistent with our theoretical assumption that the head-teacher incurs monitoring cost C^H to send a report.

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

3.2 Data Description

Our analysis draws from three sources of data: our own independent spot-checks of teacher attendance, head-teacher reports of teacher attendance submitted by mobile phone, 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 (Term 2, pre-intervention), 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 additional information about school and teacher characteristics.

Table 2 provides descriptive statistics of key school and teacher characteristics. In our sample of 85 schools, there are on average 9.28 teachers per school, leading to a total sample of 789 teachers, who are predominantly male (58 percent), 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 74 percent. This is in line with what has been observed in earlier studies on Uganda and other low income countries (Chaudhury et al. 2006, Barr, Mugisha, Serneels & Zeitlin 2012, Wane & Martin 2013).

Table 3 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) indicate 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 (6) 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 the 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.

3.3 Empirical Strategy

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

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

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 refers to district fixed effects; $(\text{Info})_s$ and $(\text{Info\&Bonus})_s$ are the two treatment dummies; and $\varepsilon_{i,s,t}$ is the error term clustered at the school level, using the wild cluster bootstrap resampling method (Cameron & Miller 2015). Our preferred specification thus pools 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.[13] When our dependent variable is teacher attendance we can also control for baseline data, since we conducted independent spot-checks

[13] In Appendix B we allow for time-varying treatment effects and fail to reject the null of coefficient equality over time.

prior to the intervention. Here the estimating equation is

$$Y_{i,s,t} = \theta Y_{i,s,PRE} + \sum_{t=1}^{3} \delta_t + \gamma_1 \left(\text{Info}\right)_s + \gamma_2 \left(\text{Info\&Bonus}\right)_s + \rho_d + \varepsilon_{i,s,t} \tag{4}$$

where $Y_{i,s,PRE}$ is baseline attendance for teacher i in school s. Finally, when we compare only the two treatment arms we estimate the following equation for a restricted sample which excludes the Control schools

$$Y_{i,s,t} = \sum_{t=1}^{3} \delta_t + \gamma_2 \left(\text{Info\&Bonus}\right)_s + \rho_d + \varepsilon_{i,s,t}. \tag{5}$$

4 Experimental Impacts

In this section we report the main impacts from the field experiment. Figure 5 presents these findings graphically, while Table 4 shows the underlying regression results.[14]

Figure 5 Panel (a) shows that teacher attendance increased when financial incentives were attached to local monitoring. On the days that we conducted independent spot-checks, teachers were 9 percentage points more likely to be present in the Info & Bonus schools compared to the Info and Control schools. Column (1) of Table 4 confirms that these differences are statistically significant. Attendance was not significantly higher in the Info schools relative to the Control schools.

Figure 5 Panel (b) shows that reported teacher attendance was also higher when the bonus scheme was available. In the Info & Bonus schools, teachers were reported as present on 61% of teacher-days with spot-checks, which is 15 percentage points higher than in the Info arm.[15] Column (3) of Table 4 confirms that this difference is statistically significant.

We next turn to policy mistakes, shown in Figure 5 Panel (c). To assess the likelihood of a policy mistake, we compare our spot-check of teacher attendance with the local monitoring report for the same teacher on the same day (if one was submitted). In the figure, the light-blue bars reflect the proportion of teacher-days with spot-checks that the teacher was absent and no report was submitted. This number is 17 percentage points higher in the Control than in the Info arm (33 percent rather than 16 percent); and 8 percentage points higher in the Info than the Info&Bonus arm (16 percent rather than 8 percent). Column (5) in Table 4 shows that these differences are statistically significant at the 1 percent level. The maroon bars reflect the proportion of teacher-days with spot-checks where an absent teacher was falsely reported as present. Contrary to the theoretical prediction, we also observed false reporting in the Info arm.[16] Furthermore, the head-teacher was no more likely to falsely report an absent teacher as present in the Info & Bonus arm. For completeness, the grey bars depict teacher-days with spot-checks where the monitor falsely reported a present teacher as absent. In both treatment arms this number is sufficiently

[14] In Table 4, Columns (1) and (3) to (5) report the results of a simple cross-sectional comparison, estimated using equations (3) and (5). Column (2) also controls for baseline presence and is thus estimated using equation (4). The first two columns show the impact of the two treatment arms relative to control, and are thus estimated using equations (3) or (4). The third columns shows the impact of the Info & Bonus arm relative to the Info arm and is thus estimated using equation (5).

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

[16] An obvious way to rationalize this finding is via some degree of transferable utility in the reputational cost term δ.

close to zero to be consistent with reporting error.

Although our theory was ambiguous as to the sign of the decision effect, we can resolve this empirically. The height of each bar in Figure 5 Panel (c) shows the overall probability of a policy mistake. This probability is close to 7 percentage points higher in Info schools than in Info & Bonus schools, a difference that is statistically significant at the 5 percent level as shown in Column (4) of Table 4.[17] Local monitoring and reporting therefore allowed for better decision-making in the Info & Bonus arm than the Info arm.

To summarize, we find that the experimental impacts are broadly consistent with the theory. Teacher attendance and reported teacher attendance are higher in the Info & Bonus schools compared to Info and Control schools; and the first source of policy mistakes −head-teachers failing to monitor and report teacher absence− is higher in the Info and Control schools than in the Info & Bonus schools. The second source of policy mistakes−head-teachers falsely reporting teacher presence− runs counter to the theory, however. Rather than finding that the bonus reduced information quality, we actually observed that head-teachers were equally likely to misreport teacher presence in both treatment arms. Taken together, these results indicate that attaching financial incentives to local monitoring reports was effective at increasing teacher attendance *and* reducing policy mistakes.

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 varied between treatment arms. For example, teachers in the Info & Bonus arms might have been more responsive to our visits, if they believed it could have an implication for their bonus payments. To test for a Hawthorne effect, in the first term of the experiment we visited some schools more than once during each round of spot-checks and randomly varied the frequency of visits: some schools received three visits, and some schools only one visit. Reassuringly, when we regress teacher attendance on the number of spot-check visits in subsequent terms, controlling for district and period fixed effects (not shown), we find no evidence of a Hawthorne effect. The number of visits does not significantly impact teacher attendance.

5 Welfare Analysis

The preceding section delivered estimates of the magnitude of the attendance and payout effects, and both the sign and magnitude of the decision effect. Although these results established that attaching financial incentives to local monitoring reports was effective at increasing teacher attendance and reducing policy mistakes, this 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 the cost of bonus payments—in our experiment, an average of USD 2,250 per school—and on the social values, $n\varepsilon^P$ and κ, placed on these outcomes.

We combine our experimental estimates of the behavioral effects with complementary data on social benefits to evaluate welfare using the bureaucracy's expected payoff as the criterion, as proposed in Section 2. We focus here on what we consider to be the more interesting welfare comparison: moving

[17] The regression includes the negligible proportion of teacher-days where a present teacher was falsely reported as absent.

from an unincentivized to an incentivized local monitoring scheme, i.e., from Info to Info & Bonus.[18] Plugging the experimental impacts into this welfare criterion in equation (2), we see that the Info & Bonus scheme improves expected welfare only if $n\varepsilon^P \cdot 0.08 - \kappa \cdot (-0.07) >$ USD $2,250$. The empirical challenge of this section is gauge whether this inequality plausibly holds.

We quantify the expected total pupil benefit from the introduction of bonus payments, $n\varepsilon^P \cdot 0.08$, in three steps. First, we calculate the impact on net enrollment, using both administrative data and data from a tracked cohort of pupils. Second, we back out gains in grade attainment implied by the enrollment figures. And third we combine data from the Uganda National Panel Survey with estimates from the literature on the causal return to schooling to calculate the increase in the net present value of future lifetime earnings due to higher grade attainment.[19]

To anticipate our key conclusion, we find that for a broad set of parameter assumptions regarding the causal return to schooling, and even under extremely conservative assumptions about program impacts on induced grade attainment, it *is* welfare-enhancing to pay for locally monitored teacher attendance. In fact, for most parameter values the net social gain from attaching bonus payments to local monitoring reports is positive even before we consider the value of reduced bureaucratic mistakes. Since the probability of making a policy mistake was estimated to be lower under Info & Bonus than Info this indicates that welfare improved even under conservative assumptions. Below we summarize the analysis underlying this conclusion, relegating a detailed discussion to Appendix C.

5.1 Enrollment Gains

The starting point for our welfare analysis is the observation that the intervention had an impact upon pupil enrollment, as well as teacher attendance. Figure 6 plots average enrollment per grade in Info and Info & Bonus schools, as reported in the endline survey. Two facts stand out. First, in both treatment arms there is a downward trend in enrollment. This is consistent with the prevailing view that school dropouts are a serious concern in Uganda.[20] Second, at each grade, average enrollment is higher in Info & Bonus schools relative to Info schools, suggesting that paying for locally monitored performance may have been more successful at averting dropouts.

Table 5 verifies that the enrollment gain (or rather reduced loss) in Info & Bonus schools is statistically significant, using two different sources of data. The results in Column (1) rely on our baseline and endline survey data and are based on the estimation strategy in equation (4), with the exception that there is now only one post-intervention period.[21] Schools in the Info & Bonus arm report on average 47 more pupils enrolled across all grades compared with Control schools (8 percent increase), and 70 more pupils

[18] The welfare comparison between the Info and Control arms is simple (ignoring the fixed costs of setting up the program). Plugging the estimates in the first row of Table 4 Columns (2) and (4) into equation (18) in Appendix A, we get $n\varepsilon^P \cdot 0.01 + \kappa \cdot 0.07 > 0$.

[19] We do also consider the possibility of using data on learning outcomes. However, for reasons discussed in Section 5.5 below, we feel that this approach is not well suited to our welfare analysis.

[20] This is inline with official records indicating that nationwide only 30 percent of pupils enrolled in grade 1 make it to grade 7 (Ministry of Education and Sports, 2014:121). In our Control schools, the number of pupils in grade 7 is on average only 40 percent of the number of grade 1 pupils.

[21] Baseline survey data is missing in two schools due to enumerator error, prompting us to use 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.

compared to Info schools (13 percent increase).

This finding is corroborated in Table 5 Column (2), which reports enrollment for a cohort of 20 pupils surveyed in 2010 as part of a previous study and representative of those enrolled in Primary 3. We tracked the enrollment outcomes of these children during our endline survey in November 2013. In the Control schools, only 34 percent of these children were still enrolled in the same school three years later. In Info & Bonus schools, the percentage of the tracked cohort still enrolled in 2013 was 14 percentage points higher than in Control schools, and 9 percentage points higher than in Info schools.

The similarity of results across the two different data-sets is reassuring and suggests that the enrollment impacts are due to the introduction of financial incentives. However, a causal impact on enrollment does not necessarily imply a causal impact on grade attainment, as grade repetition and inbound transfers from other schools are also possibilities. We turn to this below.

5.2 Moving from Enrollment to Grade Attainment

To model the impact of increased enrollment on grade attainment, we back out the portion of the enrollment gain that is due to 'averted dropouts', rather than grade repeaters, using two separate empirical strategies (see Appendix C.1). Our first approach combines our survey data on enrollment with administrative data on grade repetition. The resulting predictions for averted dropouts by grade are shown in Figure C.2. Our second approach derives annual dropout and repetition rates from differences in reduced-form dropout and repetition probabilities observed for the tracked sample surveyed in 2010 and 2013. These estimated annual rates are reported in Table C.1. Armed with these predictions, we conservatively assume that each averted dropout will progress only one additional grade before dropping out.

5.3 Moving from Grade Attainment to Earnings

We use data from the 2011/12 Uganda National Panel Survey to estimate the net present value (NPV) of future lifetime earnings for each level of grade attainment, and then calculate the increase in NPV for the pupils who gain a year of schooling, given different assumptions on the causal impact of schooling on earnings and the probability of employment (see Appendix C.2). Table C.3 brings together these steps to deliver a total per-school gain in NPV of future lifetime earnings from the introduction of the bonus scheme. This is our school-level estimate of the expected total pupil benefit in equation (2).

5.4 Cost Benefit Comparison

Table 6 shows different values for the total per-school gain in NPV of future lifetime earnings, based on four different parameter assumptions and two data sources. We present results for more and less conservative assumptions on the causal impact of an additional year of education on wages[22] and the probability of entering formal employment,[23] and also for two different data sources used to estimate

[22]A figure of 6.8 percent is the most conservative value from studies that use a quasi-experimental design to credibly estimate the causal impact of education on earnings in developing countries (Card 2001, Duflo 2001). A far larger value of 15.9 percent has been used for Uganda (Montenegro & Patrinos 2014) but the analysis is based on an ordinary-least-squared estimate using a cross-sectional data-set and so does not adequately account for selection problems. The true value plausibly falls between these two figures.

[23]To the best of our knowledge, no studies in developing countries have looked at the impact of education on gaining formal employment. As an upper bound, we use our own figure, estimated when calculating the NPV of future earnings.

enrollment gains. In all cases we assume a discount rate of 10 percent.[24] The red cells indicate parameter configurations where this gain is less than the per-school bonus cost of USD $2,250$.

Comparing the total per-school gain in NPV of future lifetime earnings with the per-school cost of bonus payments, we see that for six of the eight cells in Table 6 it is welfare-enhancing to pay for locally monitored teacher attendance, even before we consider the decision effect. Our preferred estimate is USD 2,722, which is higher than the school-level bonus cost of USD $2,250$.[25]

For the two most conservative configurations (in red) however, κ must be positive for the program to be welfare-enhancing. In the worst case the attendance and payout effects imply a net welfare loss of USD 505. Recall that our estimate of the decision effect was -0.07. To justify paying for locally monitored performance in this most conservative case, the minimum value society must place on avoiding policy mistakes is therefore USD $505/0.07 \approx$ USD $7,214$ per school or, since schools are open 180 days per year, USD 40 per school day. The Government of Uganda has stated that it aims to spend no more than UGX $150,000$ (or roughly USD 44) per school inspection (Ministry of Education, Science, Technology and Sports 2014).[26] If we take USD 44 as the Government's true valuation of information, κ, then it follows that P4LMP is welfare-enhancing even under our most conservative assumptions.

5.5 Direct Impacts on Learning Outcomes

In addition to increasing years of completed schooling, teacher attendance may also have a direct impact on student learning outcomes, as has been shown elsewhere (Duflo et al. 2012). These learning impacts may be of interest per se and, in principle, also provide a complementary approach to estimating welfare. However, in our setting there are three reasons to be cautious of this approach. First, and most prosaic, there is an issue of identification: given that our sample is school-based, the differential enrollment induced by the interventions causes bounds on estimated learning impacts to be relatively uninformative. Second, placing a contextually relevant economic value on learning gains is difficult given the paucity of studies relating these to labor-market outcomes in developing countries; studies of the relationship between cognitive skills and earnings tend to be focused on OECD economies (Hanushek, Schwerdt, Wiederhold & Woessmann 2015). Third, given a set of estimates of attainment impacts, a general problem is that these partly embody learning gains, creating a risk of double-counting those benefits. With these caveats in mind, we document briefly below what the available data show regarding impacts on learning but make no attempt to incorporate these results into our cost-benefit analysis.

We conducted independent numeracy and literacy tests at the end of the year of intervention on a

The most conservative alternative assumption is that education has no impact on the probability of entering the formal labor market.

[24] A discount rate of 10 percent was used by both Ozier (2011) and Baird et al. (2015) in their welfare analysis of a deworming program in Kenya. Similarly, Dhaliwal, Duflo, Glennerster & Tulloch (2012) provide a summary of different social discount rates used in different contexts, calculated using the social opportunity cost of capital, and also conclude that 10 percent is the appropriate rate.

[25] This is our preferred estimate because it: (i) uses the cohort data (and hence requires a weaker assumption on the impact of the interventions on transfers); (ii) invokes the more conservative assumption on the returns to education on wages; and (iii) applies the more plausible assumption of a non-zero impact of a year of education on the probability of gaining employment.

[26] It is questionable whether this target has actually been met. A 2008 audit documenting the frequency of inspections and the overall inspection budget suggests a cost closer to USD 190 per visit (Ministry of Education, Science, Technology and Sports 2010).

random sample of 40 pupils in each school, 20 in grades three and six respectively. To adjust for possible sample selection bias brought on by enrollment gains,[27] we adopt a Lee Bounds approach. To calculate the lower bound, the Lee Bound estimator drops the best-performing students from the group with the lower attrition rate (in our case, the Info & Bonus arm) such that the attrition rate in each experimental arm is equal. For the upper bound the sample selection assumption is reversed, and the estimator drops the worst-performing pupils from the group with the lower attrition rate.

Table 7 shows results from constructing Lee Bounds on learning outcomes, based on the 57 schools where (in a previous study) we also tested pupils prior to the intervention. The odd-numbered columns show results for the literacy test, and 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 schools; and the final two columns compare Info and Control schools. The first and third row indicate 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; and 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 7 that large differentials in rates of retention across treatment arms invalidate any assessment of learning outcomes. Depending on the assumption relating to sample selection, one could infer that the Info & Bonus arm had a statistically significant positive or negative impact on learning gains.[28]

6 Conclusion

Received wisdom has it that pay for locally monitored performance (P4LMP) will fail to incentivize desired behavior and will bias information for planning purposes. But simple Coasian logic suggests the reverse: P4LMP could both incentivize desired behavior *and improve* decision-making.

Responding to these observations, this paper set out to answer three related questions. Can P4LMP improve service delivery? Does P4LMP reduce or improve the quality of reported information for planning purposes? And what is the overall impact of P4LMP on social welfare? To answer them, we first developed a theoretical model that fleshed out the Coasian logic, and showed why, and for which parameters, the welfare-enhancing scenario exists. Next we used data collected during a field experiment in Ugandan primary schools to estimate the sign and magnitude of the key behavioral effects identified in the theory: teacher attendance, head-teacher monitoring/reporting, and the likelihood of bureaucratic policy mistakes due to either missing or false reports of teacher absence. Finally, using a welfare criterion suggested by the theory, we combined our experimental estimates with additional administrative data and the Uganda National Panel Survey to undertake a cost-benefit analysis of alternative intervention designs. The key takeaways from this analysis are that P4LMP can improve service delivery—in our case, teacher attendance—as well as the quality of information for planning purposes. Moreover, this dual objective can be met sufficiently cheaply to pass a cost-benefit test.

[27] In our sample, students who dropped out did indeed perform worse (although the difference is not statistically significant) in both numeracy and literacy tests sat prior to the intervention, relative to their peers who remained enrolled.

[28] Note that the range is largest in Table 7 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.

A natural question is whether the P4LMP model evaluated in this paper should be rolled out at scale. It seems clear that the high rates of teacher absenteeism observed in Uganda are at least partly due to a failure of the system to provide appropriate resources and incentives for monitoring. Districts have formal responsibility for monitoring schools but are typically under-staffed with a limited transportation budget and so find it difficult to undertake regular inspections. Across Uganda there are 87 schools per inspector and so, at most, an inspector can visit each school twice a year,[29] although in practice they see schools far less.[30] Since previous research has shown that more monitoring is associated with lower teacher absence (Muralidharan, Das, Holla & Mohpal 2014), one policy response would be for the government to transfer additional resources to districts in the hope that this translates into more school inspections. The available evidence for Uganda indicates that this may not work however: the number of inspectors has more than quadrupled since 2008 with no evidence of improvement in teacher attendance.[31]

Our results suggest a cheaper and arguably more incentive compatible alternative to district-led school inspections that makes use of cheap, readily scalable digital technology. Local monitoring and reporting by mobile phone is inexpensive to set up, simply requiring the creation of a monitoring template and central database, and costs little to run as there are no transport costs or salaries to pay to inspectors. Rather than greater investment in traditional monitoring by district officials, the main policy response suggested by this paper is further decentralization with a scaling up of *local* monitoring and incentives to exploit one of the so-called "digital dividends" discussed in the 2016 World Development Report.[32]

Looking beyond the education context, there are many settings where public sector organizations do (or could) rely on reports by local monitors that are costly to verify. It is therefore of interest to ask whether P4LMP will generally prove as cost-effective as it has in Ugandan primary education. Our starting point was the observation that failures in public service delivery can be interpreted as a breakdown of Coasian bargaining. Seen in this light, the question of transaction costs becomes paramount (Dixit 1996). Our results suggest that service delivery can be improved with P4LMP in settings where: (i) local monitoring costs are low relative to central government; (ii) the local monitor shares, at least to some small degree, the preferences of the beneficiaries; and (iii) there is a lack of transferable utility between the local monitor and agent which prohibited bargaining in the first place. When all three conditions hold, P4LMP will put transferable money on the table and could improve service delivery precisely because of (not despite) the role played by side payments.

[29] In 2016 there were 281 inspectors and 24,419 schools, each open for 36 weeks per year.

[30] Monitoring failures are not unique to Uganda. In India, for example, "poor state capacities in terms of inadequate resources and systemic infirmities contribute significantly to ineffective monitoring" (Bhatty & Saraf 2016).

[31] In 2008, there were 68 inspectors, compared to 281 in 2016 (Ministry of Education, Science, Technology and Sports 2010). Chaudhury et al (2006) estimate an absence rate of 27 percent for 2006, compared to our estimate of 26 percent in 2012 and the World Bank's Service Deliverty Indicators (www.sdindicators.org) estimate of 27 percent in 2013.

[32] Since our theory predicts that the welfare gain from P4LMP is increasing in the strength of the head-teacher's preference for teacher attendance (ε^H in the model), complimentary policy efforts to recruit and retain pro-socially motivated head-teachers could also prove important.

References

Adelman, M., Blimpo, M., Evans, D. K., Simbou, A. & Yarrow, N. (2015), 'Can information technology improve school effectiveness in Haiti? Evidence from a field experiment', Unpublished, World Bank.

Aker, J. C. & Ksoll, C. (2015), 'Call me educated: Evidence from a mobile monitoring experiment in Niger', CGD Working Paper No. 406.

Baird, S., Hicks, J. H., Kremer, M. & Miguel, E. (2015), 'Worms at work: Long-run impacts of a child health investment', NBER Working Paper 21428.

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.

Banerjee, A. V., Banerji, R., Duflo, E., Glennerster, R. & Khemani, S. (2010), 'Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India', *American Economic Journal: Economic Policy* **2**(1), 1–30.

Barr, A., Mugisha, F., Serneels, P. & Zeitlin, A. (2012), 'Information and collective action in community monitoring of schools: Field and lab experimental evidence from Uganda', Unpublished, University of Oxford.

Best, M., Brockmeyer, A., Kleven, H., Spinnewijn, J. & Waseem, M. (2015), 'Production vs revenue efficiency with limited tax capacity: Theory and evidence from Pakistan', *Journal of Political Economy* **123**(6), 1311–1355.

Bhatty, K. & Saraf, R. (2016), Does governments monitoring of schools work?, Technical report, Centre for Policy Research.

Callen, M. J., Gulzar, S., Hasanain, S. A. & Khan, M. Y. (2016), 'The political economy of public employee absence: Experimental evidence from Pakistan', NBER Working Paper No. 22340.

Cameron, A. C. & Miller, D. L. (2015), 'A practitioners guide to cluster-robust inference', *Journal of Human Resources* **50**(2), 317–372.

Campbell, D. T. (1979), 'Assessing the impact of planned social change', *Evaluation and Program Planning* **2**(1), 67–90.

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.

Chen, D., Glewwe, P., Kremer, M. & Moulin, S. (2001), 'Interim report on a preschool intervention program in kenya', Unpublished, Harvard University.

Dhaliwal, I., Duflo, E., Glennerster, R. & Tulloch, C. (2012), Comparative cost-effectiveness analysis to inform policy in developing countries, *in* P. Glewwe, ed., 'Education Policy in Developing Countries', University of Chicago Press, London, chapter 8, pp. 285–338.

Dhaliwal, I. & Hanna, R. (2014), 'Deal with the devil: The success and limitations of bureaucratic reform in India', NBER Working Paper No. 20482.

Dixit, A. K. (1996), *The Making of Economic Policy: A Transaction-cost Politics Perspective*, The MIT Press, Cambridge, MA.

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. (2013), 'Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India', *Quarterly Journal of Economics* **128**(4), 1499–1545.

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

Gerard, F. & Gonzaga, G. (2014), 'Informal labor and the efficiency cost of social programs: Evidence from 15 years of unemployment insurance in Brazil', Unpublished, Columbia University.

Hanushek, E. A., Schwerdt, G., Wiederhold, S. & Woessmann, L. (2015), 'Returns to skills around the world: Evidence from PIAAC', *European Economic Review* **73**, 103–130.

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.

Ministry of Education, Science, Technology and Sports (2010), Value for money audit report on inspection of primary schools, Technical report, Government of Uganda.

Ministry of Education, Science, Technology and Sports (2014), Inspection grant: Planning, implementing and accountability guidelines, Technical report, Government of Uganda.

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

Muralidharan, K., Das, J., Holla, A. & Mohpal, A. (2014), The fiscal cost of weak governance: Evidence from teacher absence in India, Technical report, National Bureau of Economic Research.

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

Neal, D. (2013), 'The consequences of using one assessment system to pursue two objectives', *Journal of Economic Education* **44**(4), 339–352.

Ozier, O. (2011), 'Exploiting externalities to estimate the long-term effects of early childhood deworming', Unpublished, University of California, Berkeley.

Rothstein, J. (2011), 'Review of *Learning about Teaching*', National Education Policy Center, Think Twice Think Tank Review Project.

Sandefur, J. & Glassman, A. (2015), 'The political economy of bad data: Evidence from African survey and administrative statistics', *The Journal of Development Studies* **51**(2), 116–132.

Suri, T., Jack, W. & Stoker, T. M. (2012), 'Documenting the birth of a financial economy', *Proceedings of the National Academy of Sciences* **109**(26), 10257–10262.

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.

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.

World Bank (2016), Digital dividends, Technical report, World Bank.

Table 1: Comparison of the Info and Info & Bonus arms— the 6 parameter regions

	Causal impact of introducing bonus				Parameters		
Region	Teacher Attendance	Bonus Payout	Bureaucratic Decision-making		C^T	C^H	Ex post net welfare gain
1	None	Yes	Negative	False vs. truthful report	High	Negative	$-\beta - \kappa < 0$
2	Positive	Yes	None	Both truthful	Mid	Negative	$n\epsilon^P - \beta > 0$
3	Positive	Yes	Positive	Truthful vs. no report	Mid	Mid	$n\epsilon^P - \beta + \kappa > 0$
4	None	Yes	None	Both truthful	Low	Low	$-\beta < 0$
5	None	Yes	None	False vs no report	High	High	$-\beta < 0$
6	None	No	None	Both no report	—	High	0

Note: The parameter regions are depicted graphically in Figure 4. The final column shows the ex post net welfare gain from introducing the bonus. It is calculated as the difference in the bureaucracy‘s realized payoff in the Info & Bonus arm vs. the Info arm after costs have been drawn and all players have made their moves. The combined area of Region 2 and 3 gives the attendance effect; the combined area of Regions 1 to 5 gives the payout effect; and the area of Region 1 minus 3 gives the decision effect.

Table 2: Descriptive Statistics

	Observations	Mean	Std. Dev.	Min.	Max.
Teacher characteristics					
Teacher presence	789	0.74	0.44	0	1
Proportion of female teachers	789	0.42	0.49	0	1
Average teacher age	767	35.07	8.40	19	62
School characteristics					
Teachers	85	9.28	4.17	3	22
Pupils enrolled per school	85	538.06	298.85	74	1611

Table 3: Balance Statistics

	(1)	(2)	(3)	(4)	(5)	(6
	Control	Info	Bonus	Info v Control	Bonus v Control	Bonus v Info
Teacher Presence	0.76	0.71	0.72	-0.05	-0.05	-0.00
	(0.43)	(0.45)	(0.45)	(0.07)	(0.07)	(0.09)
Female	0.42	0.35	0.46	-0.05	0.02	0.08
	(0.49)	(0.48)	(0.50)	(0.04)	(0.04)	(0.04)
Teacher Age	35.36	35.07	34.63	-0.36	-0.49	-0.13
	(8.42)	(9.04)	(7.91)	(0.89)	(0.94)	(1.06)
Pupil Enrollment	585.34	498.28	537.56	-107.75	-47.93	59.82
	(397.99)	(282.59)	(325.23)	(84.48)	(85.35)	(80.64)
Teachers per School	9.18	8.50	10.08	-0.89	0.94	1.83*
	(4.21)	(3.02)	(4.86)	(0.88)	(1.07)	(1.05)

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. For teacher characteristics standard errors are clustered at the school level, using the Wild boostrapped resampling method. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

Table 4: Results

Panel (a): *Regression Analysis*	(1)	(2)	(3)	(4)	(5)
	Teacher Attendance		Reported Attendance	Policy Mistakes	Abs, No Rep
Info	0.00283	0.00713		-0.0709*	-0.175***
	(0.05)	(0.04)		(0.04)	(0.06)
Info&Bonus	0.0860**	0.0886**	0.149**	-0.141***	-0.260***
	(0.04)	(0.04)	(0.07)	(0.05)	(0.08)
Strata indicators	Yes	Yes	Yes	Yes	Yes
Baseline Control	No	Yes	No	No	No
Obs	3288	3288	1784	3288	3288
Control mean	0.666	0.666		0.334	0.334
Info mean	0.670	0.670	0.459	0.263	0.158
Info&Bonus mean	0.759	0.759	0.612	0.191	0.074
Panel (b): *Test Theoretical Predictions*					
Info&Bonus vs Info	0.0839*	0.0811*	0.149**	-0.0753**	-0.0907**
	(0.05)	(0.05)	(0.07)	(0.04)	(0.04)
Observations	1784	1784	1784	1784	1784

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 (3) and (5). Column (2) controls for baseline attendance, based on equation (3). 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, using the wild bootsrapped resampling method. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

Table 5: Enrollment Impacts

	(1) Admin. data (school-level)	(2) Cohort data (pupil-level)
Info	-23.52	0.0469
	(26.15)	(0.06)
Info&Bonus	46.50*	0.138**
	(24.12)	(0.06)
Strata indicators	Yes	Yes
Baseline Control	Yes	No
Obs	85	1140
Info=Info&Bonus: p-value	0.016	0.272
Control=Info&Bonus: p-value	0.058	0.024
Control mean	556.300	0.344

Note: Each column reports a separate regression on treatment assignment dummies, controlling for district fixed effects. The dependent variable in Column (1) is total school enrollment, and in Column (2) is 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 (SEs) are in parentheses. SEs in column (2) are clustered at the school level, using the wild bootsrapped resampling method. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

Table 6: Average per-school difference between Info and Info & Bonus in the gain in NPV of future lifetime earnings due to treatment

		Admin. data		**Cohort data**	
Returns to education (wage)		**6.8%**	**15.9%**	**6.8%**	**15.9%**
Returns to education (empl.)	**0%**	2114	4943	1745	4081
	0.8%	3294	6122	2722	5057

Note: Cells report values based on different assumptions of the returns to education. Red cells indicate parameters for which the school-level estimate of the expected total pupil benefit from paying for locally monitored performance is less than the school-level bonus cost of USD 2,250. In these cells, Info & Bonus results in higher expected welfare than Info only if the social cost of a policy mistake, κ is sufficiently high.

Table 7: Lee Bounds

	Info & Bonus vs Control		Info & Bonus vs Info		Info vs Control	
	(1) Literacy	(2) Numeracy	(3) Literacy	(4) Numeracy	(5) Literacy	(6) 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 on treatment assignment using the Lee Bounds estimator. The odd-numbered columns report results from literacy tests 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.

Figure 1: Teacher attendance

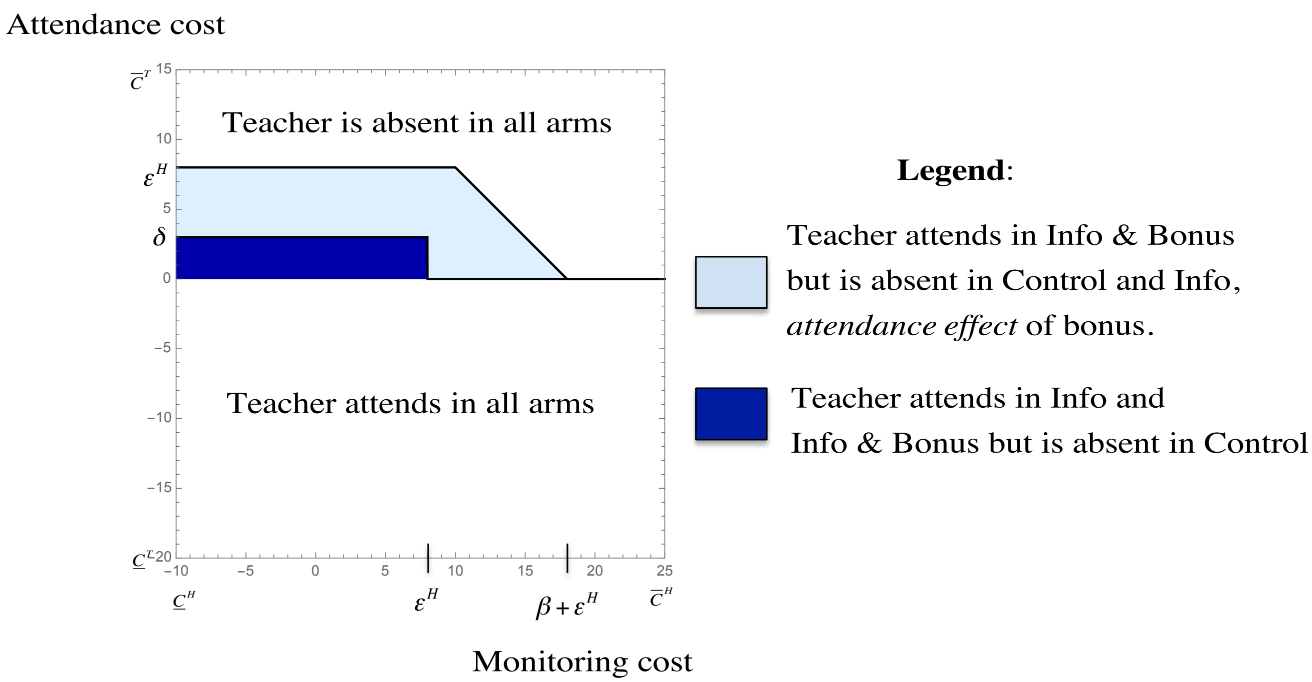


Figure 2: Head-teacher monitoring/reporting and bonus payout

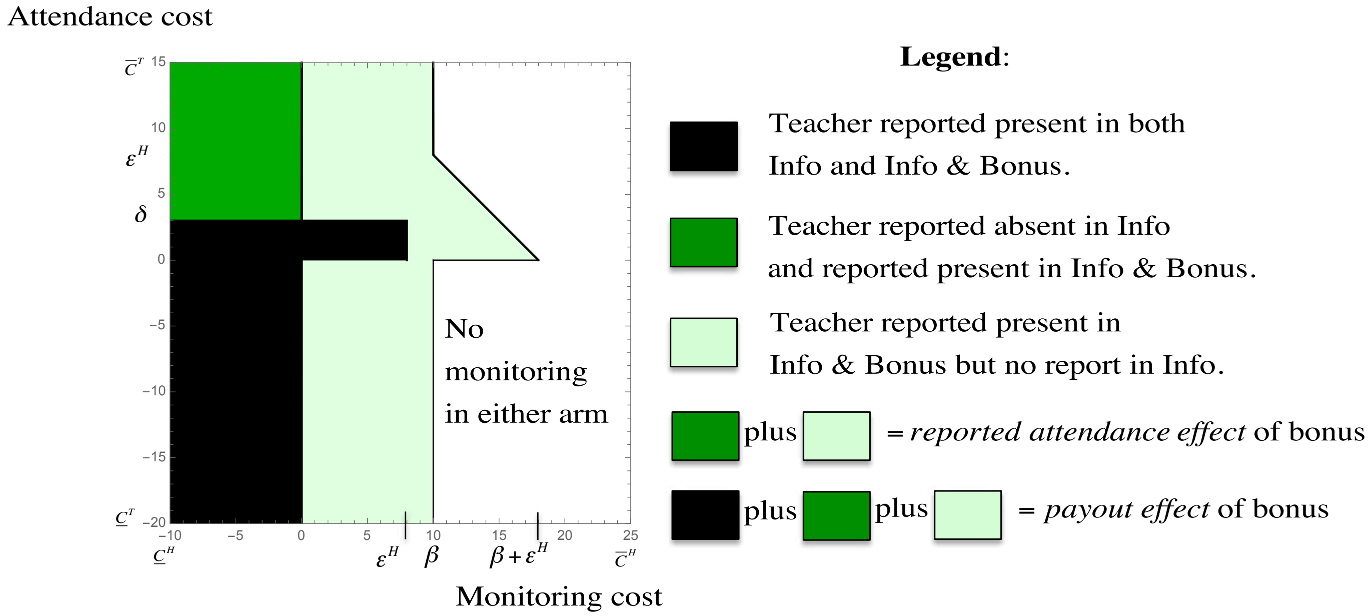


Figure 3: Bureaucratic policy mistakes

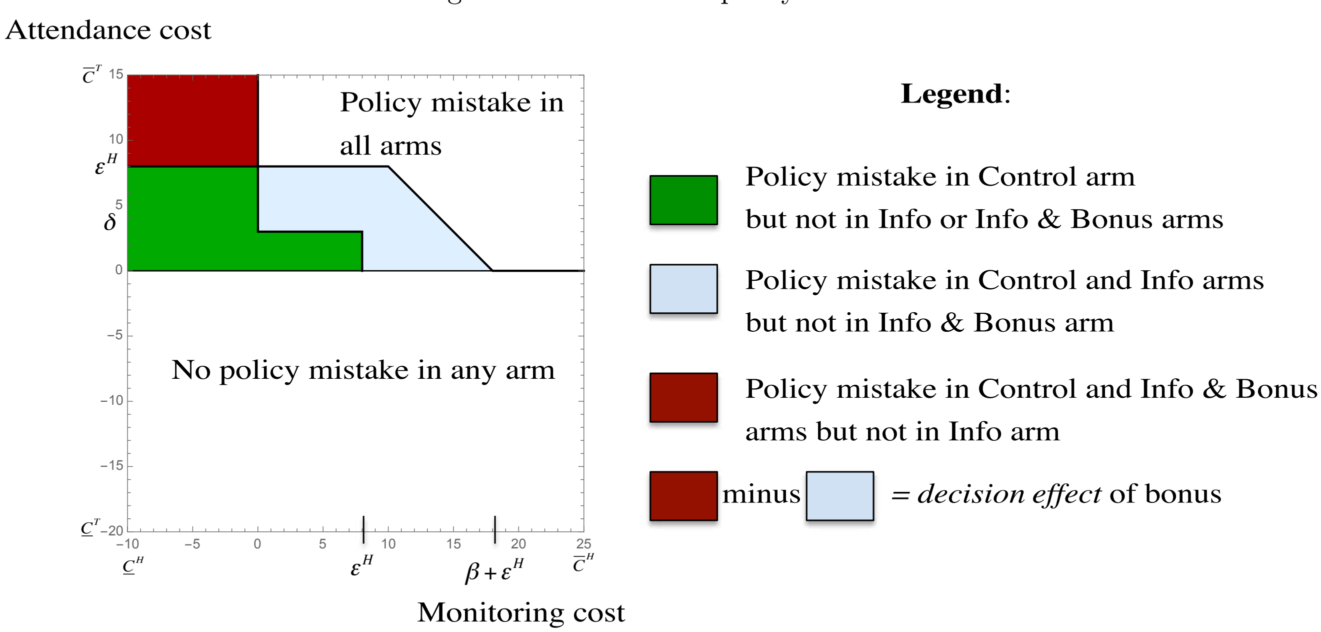


Figure 4: Comparison of the Info and Info & Bonus arms— the 6 parameter regions

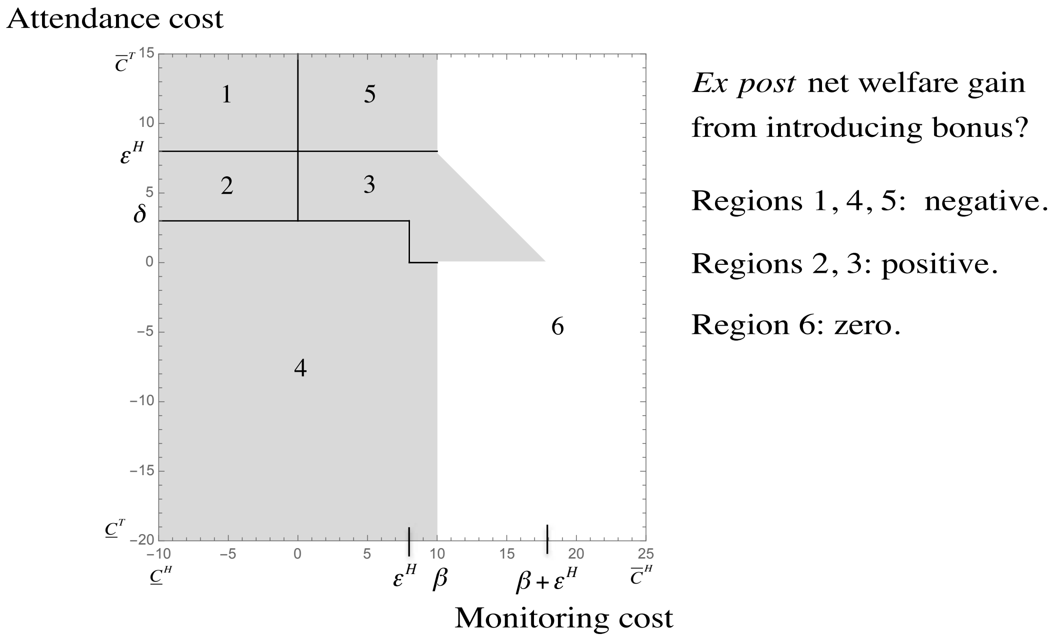
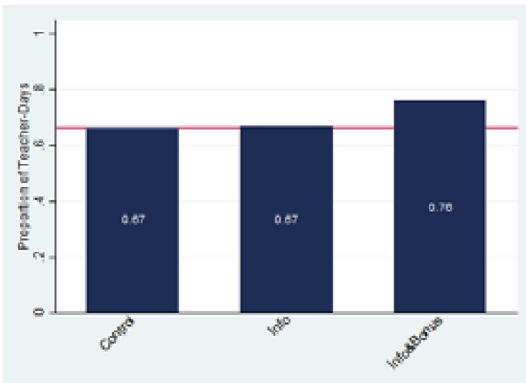
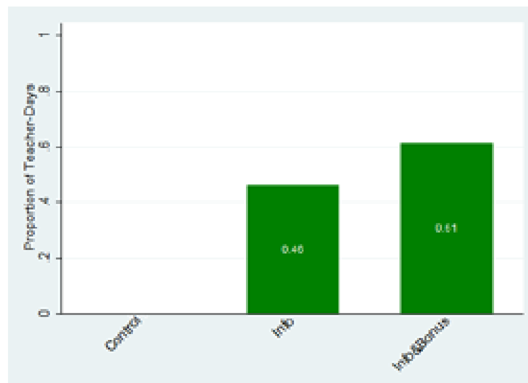


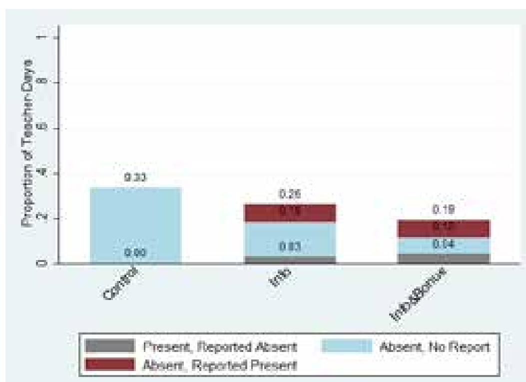
Figure 5: Experimental Impacts



(a) Teacher attendance



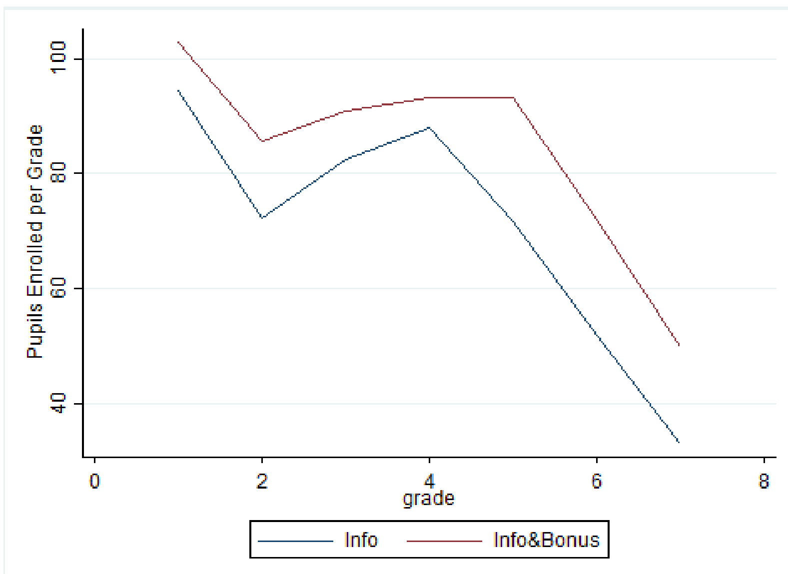
(b) Reported teacher attendance



(c) Bureaucratic policy mistakes

Note: The figure is based on 3288 teacher-days with independent spot-checks. The difference in overall bar height between Info & Bonus and Info in: panel (a) gives the attendance effect; panel (b) gives the reported attendance effect; and panel (c) gives the decision effect. The height of the Info & Bonus bar in panel (b) gives the payout effect.

Figure 6: Average enrollment at endline, by grade



Note: The figure plots the average number of pupils enrolled per grade in Info and Info & Bonus schools, as reported in the endline survey conducted in November 2013.

Appendix A Theoretical Derivations

Appendix A.1 First Best

To provide a benchmark, 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 $C^T < n\,\varepsilon^P + \delta$.

Appendix A.2 Analysis

Appendix A.2.1 Control

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

$$\Pr[a = 1|\text{Control}] = \Pr[C^T \leq 0] = \tfrac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T}. \tag{6}$$

The bureaucracy cares only about minimizing the probability of making a policy mistake. Given its 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

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

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

$$\Pr[p \neq a|\text{Control}] = \Pr[C^T > 0] = \tfrac{\overline{C}^T}{\overline{C}^T - \underline{C}^T}. \tag{7}$$

It follows that the bureaucracy's expected payoff is

$$\begin{aligned} \mathrm{E}[U^G_{\text{Control}}] &= n\,\varepsilon^P \cdot \Pr[a = 1|\text{Control}] - \kappa \cdot \Pr[p \neq a|\text{Control}] \\ &= n\,\varepsilon^P \cdot \tfrac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} - \kappa \cdot \tfrac{\overline{C}^T}{\overline{C}^T - \underline{C}^T}. \end{aligned} \tag{8}$$

Appendix A.2.2 Info

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

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

The probability of monitoring and reporting of teacher attendance is

$$\begin{aligned} \Pr[m = 1, r = 1|\text{Info}] &= \Pr[C^T \leq 0, C^H \leq 0] + \Pr[0 < C^T \leq \delta, C^H \leq \varepsilon^H] \\ &= \tfrac{-\underline{C}^T - \underline{C}^H}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)} + \tfrac{\delta(-\underline{C}^H + \varepsilon^H)}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}, \end{aligned} \tag{10}$$

and the probability of monitoring and reporting of teacher absence is

$$\begin{aligned} \Pr[m = 1, r = 0|\text{Info}] &= \Pr[C^T > \delta, C^H \leq 0] \\ &= \tfrac{(\overline{C}^T - \delta) - \underline{C}^H}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}. \end{aligned} \tag{11}$$

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

$$
\begin{aligned}
\Pr[p \neq a|\text{Info}] &= \Pr[C^T > 0] - \Pr[C^T > 0, C^H \leq 0] - \Pr[0 < C^T \leq \delta, 0 < C^H \leq \varepsilon^H] \\
&= \frac{\overline{C}^T}{\overline{C}^T - \underline{C}^T} - \frac{-\underline{C}^H \overline{C}^T + \varepsilon^H \delta}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}.
\end{aligned} \tag{12}
$$

It follows that the bureaucracy's expected payoff is

$$
\begin{aligned}
\mathrm{E}[U^G_{\text{Info}}] &= n\,\varepsilon^P \cdot \Pr[a=1|\text{Info}] - \kappa \cdot \Pr[p \neq a|\text{Info}] \\
&= n\,\varepsilon^P \cdot \left(\frac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} + \frac{\delta(-\underline{C}^H + \varepsilon^H)}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)} \right) - \kappa \cdot \left(\frac{\delta(-\underline{C}^H + \varepsilon^H) + (\overline{C}^T - \delta) - \underline{C}^H)}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)} \right).
\end{aligned} \tag{13}
$$

Appendix A.2.3 Info & Bonus

Under our parameter assumptions, 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 false 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.

- *False 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 a false 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

$$\begin{aligned}\Pr[a = 1|\text{Bonus}] &= \Pr[C^T \leq 0] + \Pr[0 < C^T \leq \varepsilon^H, C^H \leq \varepsilon^H + \beta - C^T] \\ &= \tfrac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} + \tfrac{\varepsilon^H(-\underline{C}^H + \beta) + (\varepsilon^H)2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}.\end{aligned} \tag{14}$$

The probability of monitoring and reporting of teacher attendance (equivalently the probability of a bonus payment) is

$$\begin{aligned}\Pr[m = 1, r = 1|\text{Bonus}] &= \Pr[C^H \leq \beta] + \Pr[\beta < C^H < \varepsilon^H + \beta - C^T] \\ &= \tfrac{\beta - \underline{C}^H}{\overline{C}^H - \underline{C}^H} + \tfrac{(\varepsilon^H)^2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)},\end{aligned} \tag{15}$$

and the probability of monitoring and reporting of teacher absence is zero.

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 send a false report), the bureaucracy deduces that the probability of teacher attendance at this information set is

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

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

$$\begin{aligned}
\Pr[p \ne a|\text{Bonus}] &= \Pr[C^T > 0] - \Pr[0 < C^T \le \varepsilon^H, C^H \le \varepsilon^H + \beta - C^T] \\
&= \frac{\overline{C}^T}{\overline{C}^T - \underline{C}^T} - \frac{\varepsilon^H(-\underline{C}^H + \beta) + (\varepsilon^H)^2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)}. \qquad (16)
\end{aligned}$$

It follows that the bureaucracy's expected payoff is

$$\begin{aligned}
\mathrm{E}[U^G_{\text{Bonus}}] &= n\,\varepsilon^P \cdot \Pr[a = 1|\text{Bonus}] - \beta \cdot \Pr[m = 1|\text{Bonus}] - \kappa \cdot \Pr[p \ne a|\text{Bonus}] \\
&= n\,\varepsilon^P \cdot \left(\frac{-\underline{C}^T}{\overline{C}^T - \underline{C}^T} + \frac{\varepsilon^H(-\underline{C}^H + \beta) + (\varepsilon^H)^2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)} \right) \\
&- \beta \cdot \left(\frac{\beta - \underline{C}^H}{\overline{C}^H - \underline{C}^H} + \frac{(\varepsilon^H)^2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)} \right) \\
&- \kappa \cdot \left(\frac{\overline{C}^T}{\overline{C}^T - \underline{C}^T} - \frac{\varepsilon^H(-\underline{C}^H + \beta) + (\varepsilon^H)^2/2}{(\overline{C}^T - \underline{C}^T)(\overline{C}^H - \underline{C}^H)} \right). \qquad (17)
\end{aligned}$$

Appendix A.2.4 Summing up

1. **Teacher attendance.** Highest in the Info & Bonus arm and lowest in the Control arm.

   Follows from a comparison of (6), (9), and (14).

2. **Reported teacher attendance.** Highest in the Info & Bonus arm.

   Follows from a comparison of (10) and (15).

3. **Policy mistakes.** Highest in the Control arm but an ambiguous comparison between the two treatment arms. Follows from a comparison of (7), (12), and (16). Recall there are two distinct parameter regions:

Region 1 more mistakes in the Info & Bonus arm than in the Info arm due to false reporting of teacher presence in the former and truthful reporting of teacher absence in the latter; and

Region 3 fewer mistakes in the Info & Bonus arm than in the Info arm due to truthful reporting of teacher presence in the former and no reporting of teacher absence in the latter.

Appendix A.2.5 Expected Welfare

Subtracting (8) from (13), we have

$$\begin{aligned}\mathrm{E}[U^G_{\text{Info}}] - \mathrm{E}[U^G_{\text{Control}}] = n\,\varepsilon^P \cdot (\Pr[a=1|\text{Info}] - \Pr[a=1|\text{Control}])\\ - \kappa \cdot (\Pr[p \neq a|\text{Info}] - \Pr[p \neq a|\text{Control}]) > 0. \end{aligned} \tag{18}$$

The first term on the RHS is the expected total pupil benefit from increased teacher attendance. The second term on the RHS is the expected social gain from fewer policy mistakes. Clearly, the bureaucracy prefers the Info arm to the Control. Subtracting (13) from (17), we have

$$\begin{aligned}&\mathrm{E}[U^G_{\text{Bonus}}] - \mathrm{E}[U^G_{\text{Info}}]\\ &= n\,\varepsilon^P \cdot \overbrace{(\Pr[a=1|\text{Bonus}] - \Pr[a=1|\text{Info}])}^{\text{attendance effect}} - \beta \cdot \overbrace{(\Pr[m=1, r=1|\text{Bonus}])}^{\text{payout effect}}\\ &- \kappa \cdot \underbrace{(\Pr[p \neq a|\text{Bonus}] - \Pr[p \neq a|\text{Info}])}_{\text{decision effect, ?}}. \end{aligned} \tag{19}$$

The first term on the RHS represents the expected total pupil benefit from increased teacher attendance. The second term is the expected bonus cost. The final term is the expected decision cost and is of ambiguous sign. Theoretically, the bureaucracy could gain from fewer, or lose from more, policy mistakes.

Appendix A.3 Bureaucracy's Choice of Cash Bonus

Under our parameter assumptions, if the bureaucracy chooses the Info & Bonus arm, then the optimal bonus level satisfies $\beta > \varepsilon^H$. To see why, it is helpful to consider three separate cases.

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 (17). 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.

Having considered these three cases, all that remains is to establish whether the bureaucracy prefers $\beta \geq \varepsilon^H$ to $\beta = 0$. Note that

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

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 our parameter assumptions,

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

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

Appendix B Additional Empirical Analysis

Here, we allow for a more general specification where the treatment effect on teacher attendance can vary across terms. We estimate the following equation:

$$Y_{is,t} = \sum_{t=1}^{3} \delta_t + \sum_{t=1}^{3} \gamma_{1,t} \left(\text{Info}_{is,t}\right) + \sum_{t=1}^{3} \gamma_{2,t} \left(\text{Info\&Bonus}_{is,t}\right) + \theta Y_{is,PRE} + \rho_d + \varepsilon_{is,t} \tag{20}$$

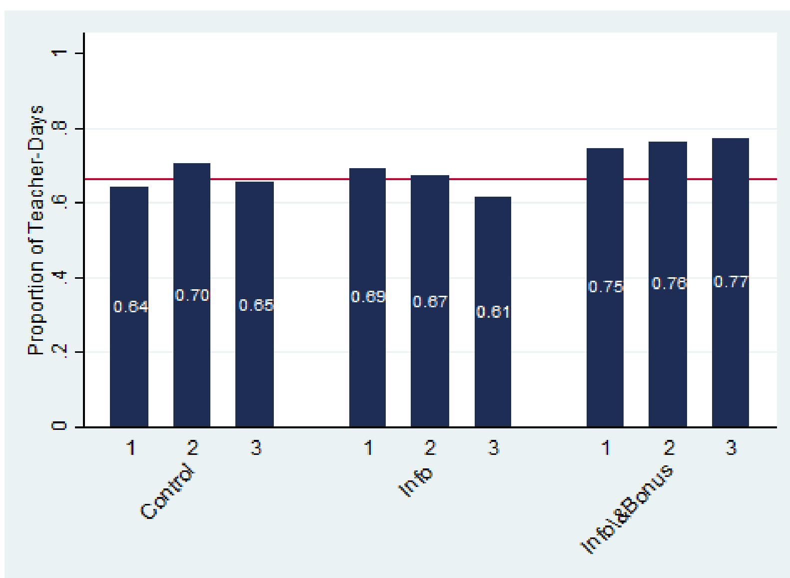
where $\text{Info}_{is,t} = \text{Info}_{is} \times \delta_t$. The treatment effects γ_{1t} and γ_{2t} are thus allowed to vary across time. We cannot reject the null hypotheses that the coefficients on each treatment arm remain the same over time; i.e. we cannot reject the null of joint equality: $\gamma_{1,1} = \gamma_{1,2} = \gamma_{1,3}$ ($p = 0.33$); and $\gamma_{2,1} = \gamma_{2,2} = \gamma_{2,3}$ ($p = 0.98$). When comparing only the Info and Info & Bonus treatment arms, we estimate:

$$Y_{is,t} = \sum_{t=1}^{3} \delta_t + \sum_{t=1}^{3} \gamma_{2,t} \left(\text{Info\&Bonus}_{is,t}\right) + \theta Y_{is,PRE} + \rho_d + \varepsilon_{is,t}. \tag{21}$$

Again, we cannot reject the null of joint equality: $\gamma_{2,1} = \gamma_{2,2} = \gamma_{2,3}$ ($p = 0.96$).

Results are also illustrated in Figure B.1, which shows teacher attendance by treatment arm and trimester of exposure. This figure shows graphically that impacts are remarkably stable across the duration of the experiment.

Figure B.1: Teacher Attendance, by term of exposure



Appendix C Detailed Welfare Analysis

This section outlines the calculations and assumptions underlying the welfare analysis presented in Table 6. We proceed in two stages. First, 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 obtain a conservative estimate of the increase in grade attainment in Info & Bonus schools relative to Info schools. Next, we use additional data sources to calculate the increase in the net present value (NPV) of future lifetime earnings caused by an increase in grade attainment.

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 grade attainment requires assumptions relating to: (i) persistence of the program and (ii) persistence of the program's impacts on attainment. In both cases we take the most conservative approach. On the first point, we model the welfare comparison for the actual experiment as it was conducted; i.e. a policy intervention for one year, with a return to the status quo and an end to project expenditures immediately thereafter. On the second point, we assume that grade attainment remains the same for all pupils, except for those who would have dropped out were it not for the program (the averted dropouts) and, furthermore, that these averted dropouts go on to drop out immediately after withdrawal of the treatment and so only gain one more year of education.[33] These conservative assumptions, again, allow us to estimate the lower bound for welfare analysis.

Estimating averted dropouts, Approach 1 Our two data sources allow for two different strategies, each with different identifying assumptions on transfers. First, using our survey data and combining it with administrative data on repetition figures in 2011 and 2012, we can back out the implied number of dropouts in grade g and year t, $\Delta_{g,t}$.

Enrollment in grade g at period t can be decomposed into the following end-states:

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

where $\Delta_{g,t}$ denotes the number of pupils who dropout at the end of the year, $\rho_{g,t}$ the number who repeat the grade, $\tau_{g,t}$ the number who transition to the next grade, and $\lambda_{g,t}$ the number who transfer out to another school. Similarly, enrollment in grade $g+1$ at the beginning of year $t+1$ can be decomposed as:

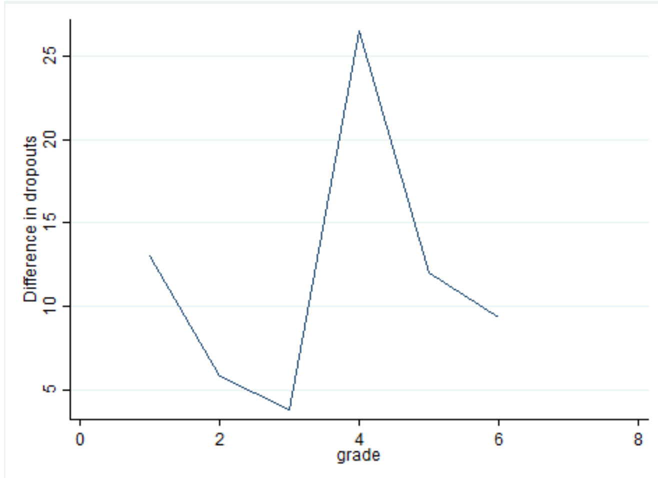
$$\pi_{g+1,t+1} \equiv \tau_{g,t} + \rho_{g+1,t} + \varphi_{g,t},$$

where $\tau_{g,t}$ denotes the number of pupils who have progressed from the previous grade, $\rho_{g+1,t}$ the number who are repeating the grade, and $\varphi_{g,t}$ the number who have transferred in from another school. Substituting in for $\tau_{g,t}$, we have:

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

[33] 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.2: Difference in dropouts between Info and Info & Bonus schools, by grade



Note: The figure shows the difference in the average number of dropouts between Info and Info & Bonus schools, by grade, where dropouts have been calculated based on equation (22) using our survey data on enrollment and administrative data on repetition. We refer to these differences as the 'averted dropouts'.

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

$$\hat{\Delta}_{g,t} = (\pi_{g,t} - \pi_{g+1,t+1}) - (\rho_{g,t} - \rho_{g+1,t})\,. \tag{22}$$

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

Figure C.2 shows the difference in dropouts between Info & Bonus and Info for each grade, estimated using (22). Note that this difference is highest in Grade 4, precisely the grade after which there is a large drop in enrollment (Figure 6 in the main body of the paper). On average 70 more pupils dropped out from Info schools relative to Info & Bonus schools and this difference is statistically significantly different from zero at the 5 percent level.[34]

Estimating averted dropouts, Approach 2 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 P3 pupils were first observed as part of a separate study, to 2013, when they were observed post-intervention.[35] Data are also available on the grade in

[34]Results from the regression analysis estimated using equation (3) are available upon request. In these regressions, the number of observations drops from 85 to 82, because the repetition number in 2012 is missing for three schools.

[35]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

Treatment, w	Dropout, $\hat{\delta}_w$	Repetition
Control	0.299	0.044
Info	0.143	0.060
Info & Bonus	0.012	0.063

which these pupils were enrolled (if any). To back out annual dropout and enrollment probabilities, we make note of the fact that three academic years were completed between the time this sample of pupils was drawn to our endline survey, but that only one of these years was spent under treatment.

Let δ_w denote the probability of dropout under treatment regime w, and recall from Table 5 that in the Control arm the fraction of the tracked cohort observed to still be enrolled at endline was 0.344. The probability of dropout in the Control arm can therefore be written as $\Pr[\text{Dropout}|w = \text{Control}] = 1 - (1 - \delta_{\text{Control}})^3 \approx 1 - 0.344$. Using the balance implied by the experimental design, we can generate observed probabilities of dropout with one period of treatment exposure in either of the treatment arms $w \in \{\text{Info, Info\&Bonus}\}$. The implied annualized dropout rates for these arms are given by setting the corresponding observed dropout probability in treatment arm w equal to $\delta_{\text{Control}} + (1 - \delta_{\text{Control}})\delta_{\text{Control}} + (1 - \delta_{\text{Control}})^2\delta_w$. The resulting implied annual transition probabilities are given in Table C.1.
Relative to the estimates of Table 5, dropout rates are lower since these represent annual rather than cumulative probabilities. Moreover, 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 *net transfers* are not different between treatment arms. On the other hand, to derive dropout rates using the tracked sample, we need to assume that *outbound transfers* are not affected by the interventions.

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 the Control and Info arms. With our conservative approach we assume that each averted dropout amounts to no more than one additional year of grade 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 Grade Attainment to Earnings

As a final step in the welfare analysis, we place a financial value on the increase in grade attainment. We begin by estimating 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 aged t currently residing in one of the 6 districts where our study took place, and who dropped out after completing grade s, using the standard Mincer

Table C.2: Predicting wages and the probability of employment

	(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) report regression results from equations (23) and (24) 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.

earnings function (and further restricting the sample to individuals who have obtained at most grade 7 and are 60 years or younger):

$$\ln w_{s,t} = \beta_0 + \beta_1 e + \beta_2 e^2 + \beta_3 s + \varepsilon_{s,t}, \tag{23}$$

where s is years of schooling and $e = t - s - 7$ is the number of years of experience.[36] Since not all of our sample have formal employment, we also estimate the expected probability, $P_{s,t}$, of earning a wage:[37]

$$P_{s,t} = \alpha_0 + \alpha_1 e + \alpha_2 e^2 + \alpha_3 e^3 + \alpha_4 s + \varepsilon_{s,t}. \tag{24}$$

Note that equations (23) and (24) need not make any causal claims on the returns to education. At this point we do not wish to estimate the impact of grade attainment on earnings; we merely want to estimate the right counter-factual—the expected net present value of future lifetime earnings for each level of primary school grade attainment. Table C.2 shows the results of these two regressions.

Using the predicted evolution of wages and employment probability over time, $\hat{w}_{s,t}$ and $\hat{P}_{s,t}$, and

[36] The data do not allow us to determine the age at which someone completed a grade. So we need to assume that everyone starts school aged 7 and progresses through grades at the same rate.

[37] We adopt a linear probability model specification (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. However, even though we use a linear probability model, the probability of formal employment always remains between zero and one.

assuming that everyone stops working at age 60 and no-one repeats a grade, we can now estimate the lifetime earnings of an individual who dropped out at the end of grade s (discounting back to the year that the individual dropped out of school and joined the labor market):

$$N\hat{P}V_s = \sum_{t=s+7}^{t=60} \left(\frac{1}{1+r}\right)^{t-s-7} \left(\hat{P}_{s,t} \cdot \hat{w}_{s,t} + \left(1 - \hat{P}_{s,t}\right) \cdot A\right). \tag{25}$$

To simplify the analysis we assume that there are only two sectors, formal wage employment and subsistence agriculture. We further conservatively assume that earnings from agriculture, A, do not depend on years of education and experience.[38] 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.[39]

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 on wages and the probability of employment. We use 6.8 percent as a lower bound of the causal impact of education on earnings since this is the most conservative estimate from studies that credibly estimated the causal impact of education on earnings in developing countries.[40] As an upper bound we use 15.9 percent, the value calculated by Montenegro and Patrinos (2014) To the best of our knowledge, no studies in developing countries have looked at the impact of education on gaining formal employment. Therefore, as an upper bound, we use our own estimate from equation (23), $\hat{\alpha}_4 = 0.008$. In Table 6, we also show results under the most conservative assumption of no impact of education on the probability of gaining formal employment.

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:

$$\frac{\partial NPV_s}{\partial s} = \sum_{t=s+7}^{t=60} \left(\frac{1}{1+r}\right)^{t-s-7} \left(0.068 \cdot \hat{w}_{s,t} \cdot \hat{P}_{s,t} + 0.008 \cdot (\hat{w}_{s,t} - A)\right). \tag{26}$$

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, i.e. comparing Info & Bonus with Info schools. Column (2) shows the expected net present value of future expected lifetime earnings, for each level of grade attainment (using equation (25), setting $r = 10$).[41] Column (3) the expected improvement in NPV due to one additional year of schooling completed (using equation 26, assuming a causal impact of 6.8 percent and 0.8 percent on wage earnings and probability of formal employment

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

[39]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).

[40]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 for Indonesia. Other well-identified studies that use samples from developing countries (e.g. Card 2001) typically find larger returns to education.

[41]A social discount rate of 10% is commonly used in the education literature (for example, Ozier (2011) and Baird et al. (2015)). Dhaliwal et al. (2012) survey different social discount rates and conclude that "10 percent is a reasonable rate for discounting. . . in developing countries".

respectively). Columns (4) and (6) show the results from Section C.1: the average reduction in dropouts per grade per school, due to the introduction of financial incentives. In Column (4), this is calculated using the administrative data on enrollment and repetition; in Column (6) this is estimated using the cohort data. Finally, Columns (5) and (7) show the average financial gain per grade per school.

Table C.3: Calculating the average per-school difference between Info and Info & Bonus in the gain in NPV of future lifetime earnings due to treatment

			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: This table gives the underlying calculations for one set of parameter assumptions, and the two data sets, reported in Table 6. Column (2) shows the NPV of future lifetime earnings, given each grade attainment, assuming a discount rate of 10%. Column (3) 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 (4) and (6) indicate treatment effects −the average number of averted dropouts per grade per school due to the program− calculated using the two different data sets. Columns (5) and (7) show the average financial gain per grade per school. The figures of USD 3,294 and USD 2,722 are reported in first and third cells in the bottom row in Table 6.

Economic Policy Research Centre
Plot 51, Pool Road, Makerere University Campus
P.O. Box 7841, Kampala, Uganda
Tel: +256-414-541023/4, Fax: +256-414-541022
Email: eprc@eprc.or.ug, Web: www.eprc.or.ug