Spring 5-18-2018

Incentives and Teacher Effort: Evidence from Lagos, Nigeria

Simileoluwa Adebajo
simiadebajo@gmail.com

Incentives and Teacher Effort: Evidence from Lagos, Nigeria

Simileoluwa Adebajo

Faculty Advisor

Professor Alessandra Cassar

Department of Economics
University of San Francisco
2130 Fulton St.
San Francisco, CA 94117

Thesis Submission for the Masters of Science Degree
in International and Development Economics

e-mail: *sradebajo@dons.usfca.edu*

May 2018

Abstract: According to recent research (Hattie, 2003), teachers contribute to around 30% of the overall variation in student achievement and success – more than any other influencing factor. This study seeks to understand how different types of incentives (monetary, near monetary and non-monetary incentives) influence the "effort" of public school teachers as perceived by the students in Lagos, Nigeria using a novel measurement tool – the teaching effectiveness survey – to measure the teachers' outcomes. Using a randomized field experiment where students evaluate the changes in their teachers' effort with a standard teaching effectiveness survey and differences in differences estimation, we show how incentives could potentially improve (or harm) teaching effectiveness. The results show that monetary incentives and near monetary incentives have no significant effect on "effort" while non-monetary incentives have a significant negative effect on the effort of teachers. This could imply that the issues underlying the current state of productivity of Public school teachers in Lagos State run deeper than remuneration or accountability. In the light of some limitations to this study, avenues for future research are multifold.

Introduction

Teaching is one of the most important professions in any society. According to the United Nations, quality teachers do not just improve test scores; they have long-term positive effects on the socioeconomic status of children 20 and 30 years down the line (UN, 2013). One question that many social scientists have tried to answer in recent times is how to improve teacher quality and performance, especially in developing countries where a larger set of challenges are faced by these professionals. This paper examines how both pecuniary and non-pecuniary incentives influence teacher effort across 10 Public Secondary Schools in Lagos Nigeria.

About 85% of the world's children live in the developing world, and these children require quality education to thrive later on in life. Student learning in developing countries is often abysmal (Hanushek & Woessmann 2008), and many agree that student learning is highly reliant on teacher effort/productivity. A 2005 Organization for Economic Cooperation and Development (OECD) Report aptly states that:

"A point of agreement among the various studies is that there are many important aspects of teacher quality that are not captured by the commonly used indicators such as qualifications, experience and tests of academic ability. The teacher characteristics that are harder to measure, but which can be vital to student learning include the ability to convey ideas in clear and convincing ways; to create effective learning environments for different types of students; to foster productive teacher-student relationships; to be enthusiastic and creative; and to work effectively with colleagues and parents."

This paper will measure these sparsely measured characteristics using the teacher effectiveness survey. This can be seen as a contribution to the literature because this tool has not been used to measure teacher productivity through econometric analysis. While some may argue against the use of this tool due to its reliance on the subjectivity of the students, in the United States this tool is regularly used to evaluate teachers because it keeps them accountable both students and other key stakeholders.

Lagos State is the most economically important state of Nigeria. Asides its economic importance, Lagos is also the most populated city in West Africa – 21 million people call the bustling metropolis their home. The state has 339 Public Junior Secondary Schools and 319 Senior Secondary Schools (LASG 2016). However enrolment rates remain low (about 40%), the percentage of overage students (older than their class average) at various levels remains high (at least 30% at any given level) (UBEC 2010) and The International Labour Organization (ILO, 2003) has even lamented that the situation of teachers in the school system in the Sub-Saharan region is so bad that it had reached "an intolerable low point". Although the State has made some efforts to improve education (such as the Lagos Eko

Secondary Education Project) with favorable outcomes, this study could help us to understand how incentives could influence the productivity of teachers across schools in the state.

Using the differences in differences technique under a randomized field experiment framework, we can estimate differences in outcomes between a treatment and control group at baseline and end line so that we can attribute any differences in outcomes to the incentives given. Past literature has shown mixed results for conditional incentive schemes for teachers – some results are positive while others show counterproductive impacts. My study tests these theories using a new tool in the context of a developing country.

Literature Review

Empirical Literature Review

Concerns about teacher productivity have led many schools in the developed world to diverge from traditional teacher compensation and encourage higher productivity through various types of incentives. In the United States, the proportion of school districts implementing performance-based incentives has increased by over 40% since 2004(Imberman, 2015). Although there is not much evidence in developing countries, some studies indicate that incentives can be effective and more cost-efficient than other measures.

One of the most important contributions to this field of study (which also served as the basis for this paper) was made by Duflo et al. (2010). The paper made use of a randomized experiment and a structural model to test whether monitoring and financial incentives would reduce teacher absence and increase learning (reduced absence and increased learning being the measures of productivity) in rural India. Teacher attendance was monitored daily through the use of time-sensitive cameras, and their salaries were linked directly with attendance. The results showed that absenteeism by teachers fell by 21% relative to the control group, and children's test scores increased by 0.17 standard deviations. They estimated a structural dynamic labor supply model and found that teachers responded strongly to the financial incentives and that this alone can explain the difference between the two groups.

According to a 2009 study by Lavy, the goal of incentives for productivity is to have two major impacts. The first is to motivate teachers to apply more "effort," which includes quantity and quality. For example, to enhance quantity, teachers might spend more time on syllabus planning or after-school tutoring. They can enhance quality by adopting innovative teaching techniques or experimenting with different teaching methods. The second objective is to attract higher quality teachers. The results revealed that students had much higher academic performance which was mediated through changes in teaching methods, enhanced after-school teaching, and increased responsiveness to students' needs.

A 2015 study by the American Institutes for Research provides evidence to support the use of student surveys in measuring teaching effectiveness. The main points revealed by the research were:

- Students have substantial everyday interaction with teachers, resultant in peculiar perceptions and assessments of teacher behaviors. Furthermore, student ratings are often consistent from year to year.
- Student evaluations are a legitimate and credible source of data.
- Student assessments are more correlated with student achievement than principal evaluations and teachers' self-assessments.
- Students demographic characteristics (e.g. expected or obtained course grade, pupil gender, GPA, subject matter) did not influence teacher scores.
- Student surveys provide expedient and detailed feedback in ways that alternative measures such as grades and test scores do not.
- Appropriate use of student feedback by teachers can result in an enhanced teaching and learning environment.

It is noteworthy to observe the dearth of evidence for non-monetary incentives, thereby emphasizing the contribution of this particular study to the field. A World Bank and HDP policy research paper by Dang and King (2016) examines how incentives, both pecuniary and non-pecuniary, influence teacher effort in Lao People's Democratic Republic. The authors estimated measures of teacher effort which had been previously overlooked in the literature such as the number of hours that teachers spend preparing for classes and teacher provision of private tutoring classes. The estimation results show that teachers increase effort in response to non-pecuniary incentives.

Another momentous contribution to this literature is a cross-country study by Woessman (2010) that utilized data for about 190,000 students across 28 OECD countries to provide evidence on the link between salary adjustments for teacher performance (at the national level) and student achievement on the 2003 Programme for International Student Assessment (PISA) test. This study is crucial because the effects of performance-related teacher pay include long-run incentive and teacher-sorting processes that may evade experimental studies but are captured in cross-country comparisons. The results showed that using teacher salary adjustments to reward great performance is significantly associated with mathematical, science, and reading achievement in the sample. Specifically, students in countries that adjust teacher salaries for exceptional performance score about 25% higher on the international math test than students in countries without teacher performance pay, after controlling extensively for student, school, and country measures.

While the study mentioned above deals with higher income countries, Masino and Nino-Zarazua (2016) review several empirical studies to determine what works to improve the quality of student learning in developing countries. 38 carefully selected studies were included for analysis in the systematic review, and the researchers classified the studies based on their drivers of change of education quality including: (1) supply-side capability interventions that operate by providing physical and human resources, and learning materials; (2) policies that through incentives seek to influence behavior and intertemporal preferences of teachers, households, and students; (3) bottom-up and top-down participatory and community management interventions, which operate through decentralization reforms, knowledge diffusion, and increased community participation in the management of education systems. The review revealed that interventions in developing countries are more effective at boosting student performance when social norms and intertemporal choices are factored into the design of policies, and when two or more drivers of change are combined.

Many researchers have explored the impact of incentives on the academic performance of students. Imberman and Lovenheim (2015) estimated the impact of monetary incentive strength on student achievement under a group-based teacher incentive pay program. The idea of grouping the teachers introduces an additional peer pressure treatment which allows teachers to hold each other accountable. Awards were based on the performances of students within a grade, school and subject by using a share of students in a grade-subject enrolled in a teacher's classes. The results show comprehensively that teachers respond to incentives when the stakes are high enough, and student achievement rises in response to stronger group incentives as a result of increases in teacher effort.

Figlio and Kenny (2006) also studied Individual Teacher Incentives and Student Performance in the USA. The results showed that test scores are higher in schools with individual financial incentives for good performance. The effect was strongest in schools that may have weaker parental monitoring mechanisms. Further, the results revealed that the relation between teacher incentives and student performance could be attributed to either better schools adopting teacher incentives or teacher incentives causing more effort from teachers. In a 2005 World Bank study, Vegas and Umansky present a simplified schema of 7 major components of an effective system to attract, retain and motivate highly qualified teachers in multiple Latin American countries. The study focuses mostly on 1) teacher incentives that impact teachers and how long they remain in the field; and, 2) incentives that affect the work teachers do in the classroom.

Further, some researchers are interested in the dichotomy between impacts of these pay for performance schemes in private versus public schools, with majority of research in this field indicating that private schools have effectively adopted these schemes with positive impacts. One particularly interesting study in this field (Ballou, 2001) showed that the schemes are more successful in private schools because the

rewards given to teachers are not trivial (and are sparsely given). Moreover, private schools do not suffer certain circumstances peculiar to the public education system, especially the opposition of teacher unions.

One of the major arguments against incentives schemes for teachers is that extrinsic incentives will displace intrinsic motivation. This phenomenon is known as motivational crowding out. It occurs when introducing an incentive for a task provokes a loss of intrinsic motivation. Several researchers have specifically explored this phenomenon in the context of education by trying to understand how teachers react to incentives (Glewwe et al. – 2010, Figlio and Winicki – 2002, Figlio and Getzler – 2002).

The most noteworthy of these studies is Jacob and Levitt (2013)'s research which investigated the prevalence and predictors of teacher cheating under incentives schemes in Chicago public schools. The authors develop an algorithm for detecting teacher cheating that combines information on unexpected test score fluctuations and suspicious patterns of answers for students in a classroom. Their results estimate that serious cases of teacher or administrator cheating on standardized tests occur in a minimum of 4 to 5 % of elementary school classrooms annually and the frequency of cheating responds strongly to relatively minor changes in incentives. Summarily, the results highlight the fact that high-powered incentive systems, especially those with bright-line rules, may induce unexpected behavioral distortions such as cheating.

Some researchers have even found negative and null impacts of incentive schemes on student outcomes. One of such is Gleizerman & Seifullah (2012) which discovered negative impacts of incentives schemes on test scores of students in the teacher incentive performance scheme known as the Teacher Advancement Program in Chicago, Illinois. Springer et al. (2012) find no significant effects on the achievement of students or the attitudes and practices of teachers in a New York City program where bonuses were awarded to teams of middle school teachers based on their joint contribution to student test score gains. In conclusion, incentives have been known to have mixed results and must be carefully calibrated to attain optimal outcomes.

Theoretical Background

Incentives are not a new idea in economics and have been incorporated into several models that explain human behavior including Principal-Agent Theory, Labor supply theory, Risk theory, etc. In this

section of my paper, I will explore the two major areas of theory that apply to this field of research and help us to understand how these theories inform our empirical findings. The first part of this section will detail theories of incentives and the second section will detail theories of labor productivity.

Overview of the Theory of Incentives

The notion of incentives in economics can be traced as far back as 1776, where Adam Smith discussed the concept in his writings on the determination of wages. Although he never actually used the term "incentives," Smith noted that in agricultural contexts, workmen were "apt to overwork themselves, and to ruin their health and constitution in a few years." if they were paid liberally for their work.

Incentives become necessary in any context when the objectives of a planner (e.g., principal, government) do not coincide with the objectives of the members of society (we will refer to them as agents). Usually, the planner cares about what the agents do and what they know, i.e., the objective function of the planner is dependent on the agents' information and behavior. Therefore, to encourage agents to behave in accordance with the planner's well-defined objectives, the planner will choose an incentive scheme with rules that will specify his behavior by his perception of the agents' actions and information.

The choice of the specific incentive scheme becomes more difficult for the planner if either some/all of the agents' payoff-relevant information is unknown or if the planner cannot observe agents' actions perfectly. The first scenario creates the problem of adverse selection, where individuals who do not plan on aligning with the objectives of the planner are thought to be complying. The second scenario, on the other hand, creates the issue of moral hazard, where the agents do not fully comply with the planner's objectives due to a lack of efficient monitoring. In the end, the planner chooses an incentive scheme which will maximize his payoff subject to the constraint that, given this scheme, agents will maximize their objective functions. What it means for the agents to maximize their objective function becomes more complex as the number of agents increases, creating a strategic game where the planner must now optimize subject to the agents being in equilibrium.

A rent extraction-efficiency tradeoff exists where the incentive scheme must elicit the agents' private information to efficiently achieve the planner's objectives. The issue also arises where the incentive scheme must incorporate full monitoring to efficiently achieve the planner's objectives. This can only be done by giving up a costly information rent to the privately informed agent or a rent to another agent for monitoring purposes. At the optimal second-best scenario, the planner trades-off his desire to attain efficiency against the costly rent given up to the agent.

Recent developments in incentive theory have also focused on loss aversion as a tool in designing incentives schemes, with some scholars even proposing that it could eliminate the rent extraction-efficiency tradeoff that traditional incentive schemes present. Loss aversion works so that instead of being rewarded for complying with the planner's objectives, agents are penalized for acting otherwise. It makes sense that agents will work to avoid the penalizations because empirical tests suggest that people commonly value a loss more than an equal gain by a factor of 2 to 4 or more (see, e.g., Kahneman et al. 1990). Evidence of such effects has come from both experimental and non-experimental settings (Samuelson and Zeckhauser 1988; Kahneman et al. 1991)

The nature of an incentive scheme greatly impacts its efficacy in incentivizing agents. Another area of interest to economists has been the efficacy of individual incentive schemes relative to group incentive schemes. Individual incentive schemes reward individual achievements while group schemes require teamwork and collaboration for the entire group to be rewarded. While both types of schemes have benefits as well as drawbacks, there is no perfect answer to which the best is.

Ladley et al. (2015) studied the impact of individual versus group rewards on work group performance and cooperation. The authors used computational social methods and Agent-Based Models to simulate work group interactions as different forms of iterated games. Group-based systems were found to outperform individual based and mixed systems, producing more cooperative behavior, the best-performing groups and individuals in most types of interaction games. Barnes et al. (2011) also studied this conflict between group versus individual incentives, postulating that trying to mix the two types of schemes puts team members in a social dilemma, leading them to focus on the individually based component. The authors found that in comparison to group-based only incentives, mixed individual/group incentives lead team members to perform faster but less accurately and focus on their own task work to the detriment of the group.

Another area of incentive theory that overlaps with our research occurs where incentives are looked at in relation to compensation (wages). Well-calibrated incentives have been found to be effective motivators of employees in a cornucopia of empirical studies; however, the nature of these incentives is a major point of disagreement amongst economists. Baker et al. (1988) was the first paper in the field of economics that looked into this relationship. According to the authors, the first issue that arises with pay for performance schemes lies in performance measurement, which can either be objective or subjective. Each of these options has pros and cons, but both are still used to date. Promotion-Based incentives must also be compared to Bonus-Based incentives because while Promotions are less costly than bonuses, they are usually not as effective.

Theories of Labor Productivity

Microeconomic theory implies that wages are closely related to the marginal productivity of labor. Specifically, theory suggests a clear relationship between productivity, wages and labor demand, in which wages correspond to the marginal productivity of labor and can be derived from the profit-maximizing behavior of firms. In reality, however, wages are not exactly equal to productivity, with historical data showing a wage-productivity gap that has widened progressively from the mid-1990's to date. The development of this gap has been accompanied by important changes in the way businesses compensate their employees, including the use of bonuses, incentives, etc.

Along the same vein, economists have extensively explored the impact of this wage-productivity gap and wage inequality on the productivity of labor. One of the most comprehensive investigations of this relationship was done by Policardo et al. (2014), with the findings revealing that wage inequality measured by a Gini index has a negative effect "ceteris paribus" on a country's labor productivity. That is, more wage inequality implies less labor productivity. It is important to note that the countries used in this study were all OECD countries, and so the results may differ in different contexts.

An important development of the simple neoclassical model incorporates a concept known as efficiency wages, which rejects the idea that wages are associated with the marginal productivity of workers. In contrast, this model argues that paying higher-than-market wages can be a rational choice for firms, e.g., to increase the work effort of employees. In the world set out in simple efficiency wage models, more than one type of worker exists. The workers are classified as:

- 'productive workers' whose utility is based (positively) on wages, but is also linked (negatively) to their work effort and to the likelihood of losing their job (and becoming reliant on benefits);
- 'shirking workers' who earn the same wage with less work effort, but who face a higher risk of job loss if their employer discovers the 'shirking,' and therefore a higher risk of reliance on unemployment benefits, with lower utility.

This model leads to an equilibrium where wages are set at the level at which workers decide not to 'shirk.' This implies that wages will be higher: a) the higher the effort the firms try to extract from workers; b) the higher the alternative utility from, e.g., unemployment benefits; and c) the smaller the probability that firms discover the shirking.

The most famous shirking efficiency wage model is Shapiro and Stiglitz (1984) where the shirking of the worker is not easy to monitor and shirking is easier than putting effort in. Thus, firms claim to decrease the probability of shirking by paying higher wages than the market clearing wages as the worker will lose a higher paid job. The intuition is that if all firms set the same wage and there is no unemployment, the worker simply puts no effort in at all, and if he/she is caught when shirking, the worker will get the same wage in another job, and there is no risk of unemployment. By paying a higher wage, losing a job

may end the worker in an unemployment status or a lower wage job. The risk of unemployment, the higher wage in the current job and low unemployment benefit all decrease the outside opportunity for the worker.

According to Shapiro and Stiglitz, the wage at which the worker will choose not to shirk will be when the present value of the utility from not shirking is greater than the utility from not shirking and can be seen in the formula below:

$$w \geq rV_U + (r + b + q)\frac{e}{q} \equiv \hat{w}$$

The equation implies a direct positive relationship between the discount rate, r; the required effort level, e; the utility gained from being unemployed; V_U, and the exogenous quitting rate, b, such that as these increase, the wage needed for workers not to shirk will also increase. We also see an inverse relationship between the optimal wage and the probability of being caught, q, so that as the probability of being caught increases, the wage needed to discourage shirking will decrease.

Furthermore, the productivity of labor in the public sector is often seen as being a polar opposite of productivity in the private sector, especially in under developed countries. Public sector employees are often disengaged from their jobs, with the problem manifesting in two main ways. First, disengaged employees may be less committed to their work and less productive as a result. Second, employees who lack clarity about organizational and personal goals, lack autonomy or do not believe that they are heard, are less likely to become active partners in redesigning their jobs for higher productivity working. (Pwc & Demos, 2014)

Baumol (1966) hypothesized that the technology of the public sector is labor-intensive relative to that of the private sector because the type of production undertaken leaves little scope for increases in productivity and that makes it difficult to substitute capital for labor. For example, hospitals need a minimum number of nurses and doctors per patient, and maximum class sizes place lower limits on teacher numbers in schools.

Regardless of the framework and underlying assumptions that may be used to evaluate productivity, all economists can agree that productivity and compensation (wages) are intrinsically linked so that one informs the other and vice versa. It is therefore essential for us to understand how we can manipulate wages to directly drive up productivity and indirectly impact economic growth in the long run.

Method & Data

The main hypotheses being tested are as follows:

H_0- Incentives have no significant impact on teacher effort among public secondary school teachers in Lagos State

H_1- Incentives have a significant impact on teacher effort among public secondary school teachers in Lagos State

This research made use of primary data collected from 10 Public Secondary Schools in Lagos State in the final quarter of the 2016/2017 school year. Surveyors made use of a Teacher effectiveness survey with 43 questions (all scaled from 1 to 6 with 1 being the lowest value and six the highest) which allowed students to evaluate their teachers' "effort" under seven broad categories:

1. Presentation of Content
2. Clarity of Expectations or Directions
3. Helpfulness/Availability
4. Usefulness/Clarity of Feedback on Performance
5. Encouragement of Participation/Discussion
6. Motivation
7. Overall Teaching Effectiveness

Simple additive indices are created to compound the questions in the survey into each of these broad headers to avoid having too many dependent variables by adding together the responses for relevant questions and multiplying by their mean. A final overall index variable was also created which compounds all 43 questions in the survey into one score. This overall index will be referred to as the 'Final Score.'

Lagos State has six educational districts in total. Two districts were randomly selected from the pool (Eti-Osa and Lagos Mainland), and our six treatment schools were then randomly selected from the Eti Osa district, and four control schools were chosen from the other. This was done to prevent significant spillover effects even within the districts. The survey was administered at baseline and end line (a 4-week window) in the randomly selected treatment and control schools to enable us to isolate the impact of our incentive treatment on the outcomes of interest. 6 of the schools in our sample were treatment schools that either received monetary, near-monetary or non-monetary rewards. The other four schools were used as our counterfactual control schools. Our treatment variable is simply an indicator variable where 1 indicates treated schools and 0 indicates the control schools.

Figure A: Map of Lagos State (sample areas are circled)

	Treatment – Eti Osa			Control- Lagos Mainland (1170 students)
	Monetary Treatment (750 students)	Near Monetary Treatment (743 students)	Non-Monetary Treatment (719 students)	
School	• Ireti Junior High School • Ireti Senior High School	• Falomo Junior High School • Falomo Senior High School	• Wahab Folawiyo Junior High School • Wahab Folawiyo Senior High School	• Aje Junior High School • Mobolaji Bank Anthony Junior High School • Iponri Estate Junior High School • Iponri Estate Senior High School

The incentives were awarded to the teachers with the highest average improvement in effort/productivity (as measured by the above key indicators). The monetary incentive was a cash prize of NGN 5,000 (US $15) for each teacher. There was also a "near money" treatment arm which received gift cards for a popular supermarket in Lagos equal in nominal value to the monetary reward. The non-monetary incentive was an award ceremony (with certificates and trophies) for the selected teachers in the school at a school-wide assembly.

One drawback of this data set lies in the fact that we were unable to gather demographic data (such as age, family income, sex, etc.) from the 162 teachers in our sample to ascertain whether the treatment

and control groups are adequately balanced. We are simply analyzing this data under the assumption that the two groups are balanced because of the random selection of schools in our sample from all Public Secondary Schools across Lagos State which should have students and teachers that are similar in characteristics across the board. The lack of demographic information also means that we will be unable to statistically prove the parallel trends assumption which is a major underlying assumption of difference in difference estimation.

Due to the lack of demographic data, I will provide some brief stylized facts that will enable us to compare the treatment and control schools and their local government areas to some extent:

- There are 20-25 full-time teachers in each of the schools in our sample, including both treatment and control schools.
- According to a School Census by the Lagos State Government, the teacher-pupil ratios in the districts where our study was carried out are 38:1 and 35:1 for Junior Secondary schools and 32:1 and 23:1 for Senior Secondary Schools respectively.
- Regarding school enrolment by gender:

	Eti-Osa District	Lagos Mainland District
Junior Secondary Schools	51.9%female, 48.1% male,	58.4% female, 41.6% male,
Senior Secondary Schools	51.1% female, 48.9% male,	52.6% female, 47.4% male,

- Regarding the demographic characteristics of the local government areas:

	Eti-Osa LGA	Lagos Mainland LGA
Population	390,800	449,900
Population (by gender)	56% male, 44% female,	52% male, 48% female
Age Distribution	15-64 yrs 72%, 0-14yrs 26%, 65+years 2%	15-64 yrs 67%, 0-14yrs 30%, 65+years 3%
Average Estimated Living Costs for a family of 2+2	US$184	US$144

We can, therefore, say that the schools should be fairly well balanced in terms of characteristics.

Model Specification

The key research idea we attempt to investigate is the impact of incentives on teacher effort among public secondary school teachers in Lagos State. In order to isolate the pure impact of our treatment, we make use of the Difference in Difference (DID) methodology. This method calculates the effect of a treatment (i.e., an explanatory variable or an independent variable) on an outcome (i.e., a response variable or dependent variable) by comparing the average change over time in the outcome

variable for the treatment group, compared to the average change over time for the control group. In addition to the basic linear regression assumptions, the major underlying assumption of this method is the parallel trends assumption which states that the trend of the outcome variable in the counterfactual (or control group) must be similar to its trend in the treatment group in the absence of the treatment. The basic empirical model we will be estimating is:

$$TeacherScore_{it} = \beta_0 + +\beta_1 Treatment_{it} * time + \rho Treatment_{it} + \gamma time + \mu_i$$

Where $TeacherScore_{it}$ is the performance of teacher i at time t as measured by our key indicators, $Treatment_{it}$ is an indicator variable which shows the treatment status of teacher i at time t (1 if you are in the treatment group and 0 otherwise), $time$ is an indicator variable which is 1 at the end line and 0 at baseline and β_1 tells us the differential impact on teacher effort of the assigned incentive relative to the control group. Furthermore, the standard errors in this model are clustered at the school level.

Results

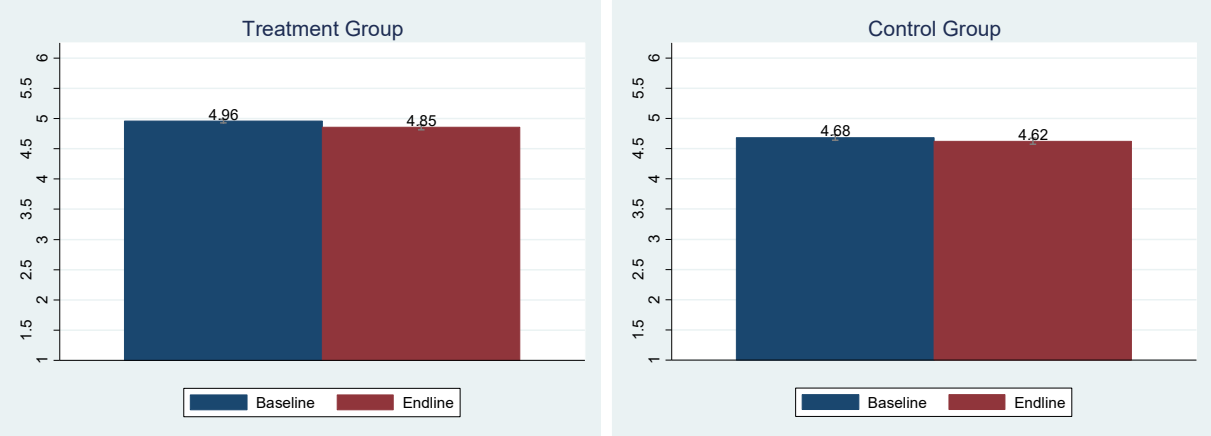

Figure 1: Baseline & Endline Mean of Final Score in Treatment (Left) and Control (Right) Groups.

Figure 1 above shows the mean of the Final Score in both the treatment and control groups at baseline and end line across the entire sample. The figure implies an overall reduction in the mean of the Final score over the course of the study across all treatment arms. Figures 2 through 4 of the appendix show similar means for the three particular incentive groups, and we can see that the near monetary and non-monetary responses seem to be driving this reduction in the mean. In all of the scenarios, the treatment

mean is significantly higher than the control means at the end line, but the difference between the groups varies by the type of incentive.

Tables 1 through 4 in the Appendix shows us the all of the Difference in Difference estimates for the entire sample of the study. The results show that overall; the incentives have no significant impact on teachers in the treatment schools relative to those in the control schools. The incentives do not have significant effects on presentation of content, clarity of expectations, teacher's helpfulness, usefulness of feedback, encouragement of participation, motivation, overall teaching effectiveness and the final evaluation score.

Following the overall analysis, we must now individually look at the impacts of the three different types of treatments relative to the control arm. Table 2 shows us the impacts of our monetary incentives on the teacher's performance. Recall that the monetary incentive was a total of NGN 5,000 (US $15) for each teacher. According to the results, monetary incentives have no significant impacts whatsoever on any of our key performance indicators including presentation of content, clarity of expectations, teacher's helpfulness, usefulness of feedback, encouragement of participation, motivation, overall teaching effectiveness and the final evaluation score.

Table 3 shows us the impact of our near monetary incentives on the teacher's performance. Recall that the near monetary incentive was a gift card for a popular supermarket in Lagos equal in value to the monetary reward. The results here also showed no significant effects of this near monetary incentive on any of our measures of teacher effort including the overarching metric – the final evaluation score.

Finally, we look at the impact of the non-monetary incentives on the performance of teachers in treatment schools relative to their peers in control schools. Recall that the reward here was an award ceremony (with certificates and trophies) for the teachers in the school at a school-wide assembly. These results in Table 4 reveal that the non-monetary incentive had highly significant negative impacts on the usefulness of feedback, encouragement of participation, motivation and the final evaluation score. This result is particularly interesting because it implies that non-monetary incentives of this nature actually hamper teaching effectiveness – possibly by fostering unhealthy competition among the teachers to win awards.

Conclusion & Recommendations

In this paper, we are able to show that incentives do not always yield desired impacts. Overall, the incentives actually do more harm than good in our sample. Although monetary incentives show positive impacts, the effects are not statistically significant. Further, the near monetary and non-monetary incentives show counterproductive results, negatively impacting the performance of the teachers in those groups.

However, these results are not completely surprising. There are many dissenting voices in economics who propose that merit-based pay for teachers will most likely do more harm than good. Lewis and Podgursky (2000) outline seven major issues that could arise as a result of merit-based pay schemes for teachers:

1. Performance-based compensation programs encourage competition rather than collaboration among teachers.
2. The Union Environment and the Collaborative Nature of Teaching are violated when some teachers are paid more than others.
3. There is no clear definition of what constitutes a "good teacher." In other words, what is merit based upon?
4. If student learning is the sole basis of the merit evaluation, relying on test scores can present major problems because teachers may be improving in other ways. (hence, my use of the teaching effectiveness survey)
5. When you reward teachers for student achievement, nobody wants to teach certain kids in certain communities, i.e., low-income children in low-income communities.
6. Bias and Favoritism between teachers and students, and even between teachers and school administration.
7. Performance-based compensation will take from teachers the ability to teach as they wish and as they do best. It just requires teachers to jump through hoops. It will make everyone teach and behave in the same way.

In the light of a few limitations to the study, we must also evaluate what could have been done better to improve the probability of detecting a significant effect. One major issue that is worth discussing is the limited period within which the study was carried out. Due to changes in the Lagos State school calendar, the entire study was done in 4 weeks from baseline to end line. Perhaps a longer period would have been able to produce more desirable results. The size/value of the incentive has also seemed to have significant impacts on outcomes in past studies. It is therefore possible that increasing the value of the incentives could have produced different results.

In conclusion, this study has revealed that trying to incentivize people to improve their performance does not always work, especially in developing contexts. It is most likely that the issues that cause poor performance amongst government teachers in developing countries run much deeper than incentives; institutions and attitudes must also be reformed to experience significant changes in behavior. Aggressive monitoring mechanisms could also be used to complement incentives schemes (as in the Duflo paper) to increase the impact of these programs. This paper contributes to a small but growing

literature that exploits structural modeling and carefully controlled randomized experiments to answer questions about education in the developing world. Avenues for further study exist in terms of different geographical contexts, various measurement tools and methodological variations.

References

Quality of teaching really matters | UN DESA | United Nations Department of Economic and Social Affairs. (2013). Un.org. Retrieved 12 May 2017, from https://www.un.org/development/desa/en/news/nocat-uncategorized/teachin.html

Hanushek, E., & Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. Journal Of Economic Literature, 46(3), 607-668. http://dx.doi.org/10.1257/jel.46.3.607

Organisation for Economic Co-operation and Development. (2005). Teachers Matter :Attracting, Developing and Retaining Effective Teachers. Paris France: Organisation for Economic Co-operation and Development. Retrieved from https://www.oecd.org/edu/school/34990905.pdf

Lagos State. (2017). En.wikipedia.org. Retrieved 12 May 2017: https://en.wikipedia.org/wiki/Lagos_State

Ministry of Education - Lagos State Government. (2017). Ministry of Education - Lagos State Government. Retrieved 12 May 2017, from http://education.lagosstate.gov.ng/

Lagos State Ministry of Education. (2010). Lagos.

International Labour Organization (2003), New York.

Imberman, S. How effective are financial incentives for teachers?. IZA World of Labor 2015: 158 doi: 10.15185/izawol.158

Duflo, E., Hanna, R., & Ryan, S. (2012). Incentives Work: Getting Teachers to Come to School. American Economic Review, 102(4), 1241-1278. http://dx.doi.org/10.1257/aer.102.4.1241

Lavy, V. (2009). Performance Pay and Teachers' Effort, Productivity, and Grading Ethics. American Economic Review, 99(5), 1979-2011. http://dx.doi.org/10.1257/aer.99.5.1979

Dang, H., & King, E. (2016). Incentives and teacher effort. Economics Of Transition, 24(4), 621-660. http://dx.doi.org/10.1111/ecot.12101

Holte, J., Kjaer, T., Abelsen, B., & Olsen, J. (2015). The impact of pecuniary and non-pecuniary incentives for attracting young doctors to rural general practice. Social Science & Medicine, 128, 1-9. http://dx.doi.org/10.1016/j.socscimed.2014.12.022

ADEEB ABDULLAH, A., & LAI WAN, H. (2013). Relationships of Non-Monetary Incentives, Job Satisfaction and Employee Job Performance. International Review Of Management And Business Research, 2(4), 105-1091.

Imberman, S., & Lovenheim, M. (2015). Incentive Strength and Teacher Productivity: Evidence from a Group-Based Teacher Incentive Pay System. Review Of Economics And Statistics, 97(2), 364-386. http://dx.doi.org/10.1162/rest_a_00486

Figlio, D., & Kenny, L. (2006). Individual Teacher Incentives And Student Performance. NBER Working Papers. http://dx.doi.org/10.3386/w12627

Vegas, Emiliana; Umansky, Ilana. 2005. Improving Teaching and Learning through Effective Incentives : What Can We Learn from Education Reforms in Latin America?. Washington, DC. © World Bank. https://openknowledge.worldbank.org/handle/10986/8694 License: CC BY 3.0 IGO.

Stark, P. (2017). Do student evaluations measure teaching effectiveness?. [online] The Berkeley Blog. Available at: http://blogs.berkeley.edu/2013/10/14/do-student-evaluations-measure-teaching-effectiveness/ [Accessed 7 Dec. 2017].

Lewis, S. and Podgursky, M. (2000). The Pros and Cons of Performance-Based Compensation. Milken Family Foundation.

Glazerman, S. and Seifullah, A. (2012). An Evaluation of the Chicago Teacher Advancement Program (Chicago TAP) after Four Years. Final Report. Mathematica Policy Research.

BAKER, G., JENSEN, M., & MURPHY, K. (1988). Compensation and Incentives: Practice vs. Theory. The Journal Of Finance, 43(3), 593-616. http://dx.doi.org/10.1111/j.1540-6261.1988.tb04593.x

Barnes, C., Hollenbeck, J., Jundt, D., DeRue, D., & Harmon, S. (2010). Mixing Individual Incentives and Group Incentives: Best of Both Worlds or Social Dilemma?. Journal Of Management, 37(6), 1611-1635. http://dx.doi.org/10.1177/0149206309360845

Baumol, William; William Bowen (1966). Performing Arts, The Economic Dilemma: a study of problems common to theater, opera, music, and dance. New York: Twentieth Century Fund.

CENTRAL BANK OF THE REPUBLIC OF TURKEY. (2011). Productivity and Wage Differentials between Private and Public Sector in the Developing Countries. Ankara: CENTRAL BANK OF THE REPUBLIC OF TURKEY.

Franz W (1999), Arbeitsmarktökonomik, 4. Auflage, Springer

Hindriks, J., & Myles, G. (2013). Intermediate public economics (pp. 73-98). Cambridge, MA: MIT Press.

Institute for Employment Studies. (2011). Wages, productivity and employment: A review of theory and international data. Brighton, UK: European Employment Observatory.

International Labour Organization (2003), New York.

Kahneman, D., Knetsch, J., & Thaler, R. (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. Journal of Political Economy, 98(6), 1325-1348. Retrieved from http://www.jstor.org/stable/2937761

Kahneman, D., Knetsch, J., & Thaler, R. (1991). Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. Journal Of Economic Perspectives, 5(1), 193-206.

Ladley, D., Wilkinson, I., & Young, L. (2015). The impact of individual versus group rewards on work group performance and cooperation: A computational social science approach. Journal Of Business Research, 68(11), 2412-2425. http://dx.doi.org/10.1016/j.jbusres.2015.02.020

Policardo, L., Punzo, L., & Sanchez Carrera, E. (2014). Wage inequality and its effects on labor productivity. Https://Editorialexpress.Com/.

Productivity in the public sector: What makes a good job?. (2014). London, United Kingdom.

Samuelson, W. & Zeckhauser, R. J Risk Uncertainty (1988) 1: 7. https://doi.org/10.1007/BF00055564

Shapiro, C., & Stiglitz, J. (1984). Equilibrium Unemployment as a Worker Discipline Device. The American Economic Review, 74(3), 433-444. Retrieved from http://www.jstor.org/stable/1804018

Smith, Adam, 1723-1790. (2000). The wealth of nations / Adam Smith ; introduction by Robert Reich ; edited, with notes, marginal summary, and enlarged index by Edwin Cannan. New York :Modern Library,

Glewwe, P., Ilias, N., & Kremer, M. (2010). Teacher Incentives. *American Economic Journal: Applied Economics*, *2*(3), 205-227. http://dx.doi.org/10.1257/app.2.3.205

Figlio, D., & Winicki, J. (2005). Food for thought: the effects of school accountability plans on school nutrition. Journal Of Public Economics, 89(2-3), 381-394. http://dx.doi.org/10.1016/j.jpubeco.2003.10.007

Figlio, D., & Getzler, L. (2003). Accountability, Ability and Disability: Gaming the System?. *Advances In Applied Microeconomics*, 35-49. http://dx.doi.org/10.1016/s0278-0984(06)14002-x

Jacob, B., & Levitt, S. (2003). Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating. *The Quarterly Journal Of Economics*, *118*(3), 843-877. http://dx.doi.org/10.1162/00335530360698441

Kreps, D. (1997). Intrinsic Motivation and Extrinsic Incentives. *The American Economic Review*, *87*(2), 359-364. Retrieved from http://www.jstor.org/stable/2950946

Fehr, E., & Goette, L. (2007). Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment. *The American Economic Review*, *97*(1), 298-317. Retrieved from http://www.jstor.org/stable/30034396

Hattie, J.A.C. (2003, October). Teachers make a difference: What is the research evidence? Paper presented at the Building Teacher Quality: What does the research tell us ACER Research Conference, Melbourne, Australia. Retrieved from http://research.acer.edu.au/research_conference_2003/4/

Kristen Underhill, When Extrinsic Incentives Displace Intrinsic Motivation: Designing Legal Carrots and Sticks to Confront the Challenge of Motivational Crowding-Out, 33 Yale J. on Reg. (2016). Available at: http://digitalcommons.law.yale.edu/yjreg/vol33/iss1/5

American Institutes for Research (2015). Uncommon Measures: Student Surveys and Their Use in Measuring Teaching Effectiveness. [online] Washington DC: American Institutes for Research. Available at: https://www.air.org/sites/default/files/Uncommon-Measures-Student-Surveys-Guidance-Nov-2015.pdf [Accessed 27 Mar. 2018].

Lagos State Ministry of Education (2010). Lagos State School Census Report 2009-2010. ANNUAL SCHOOL CENSUS REPORT. Lagos, Nigeria.: MINISTRY OF EDUCATION LAGOS STATE - FEDERAL REPUBLIC OF NIGERIA.

Masino, S., & Niño-Zarazúa, M. (2016). What works to improve the quality of student learning in developing countries? International Journal of Educational Development, 48, 53-65. http://dx.doi.org/10.1016/j.ijedudev.2015.11.012

Ballou, D. (2001). Pay for performance in public and private schools. Economics of Education Review, 20(1), 51-61. http://dx.doi.org/10.1016/s0272-7757(99)00060-6

Springer, M., Pane, J., Le, V., McCaffrey, D., Burns, S., Hamilton, L., & Stecher, B. (2012). Team Pay for Performance. Educational Evaluation And Policy Analysis, 34(4), 367-390. http://dx.doi.org/10.3102/0162373712439094

The Institute for the Study of Labor (IZA). (2010). Cross-Country Evidence on Teacher Performance Pay. Bonn, Germany: IZA. Retrieved from http://ftp.iza.org/dp5101.pdf

Eti-Osa (Local Government Area, Nigeria) - Population Statistics, Charts, Map and Location. (2018). *Citypopulation.info.* Retrieved 28 April 2018, from http://www.citypopulation.info/php/nigeria-admin.php?adm2id=NGA025008

Lagos Mainland (Local Government Area, Nigeria) - Population Statistics, Charts, Map and Location. (2018). *Citypopulation.de.* Retrieved 28 April 2018, from https://www.citypopulation.de/php/nigeria-admin.php?adm2id=NGA025015

Guzi M., Kabina. K, Tijdens KG, (2016). Living Wages in Nigeria – Lagos State. Amsterdam, Wage Indicator Foundation

Appendix

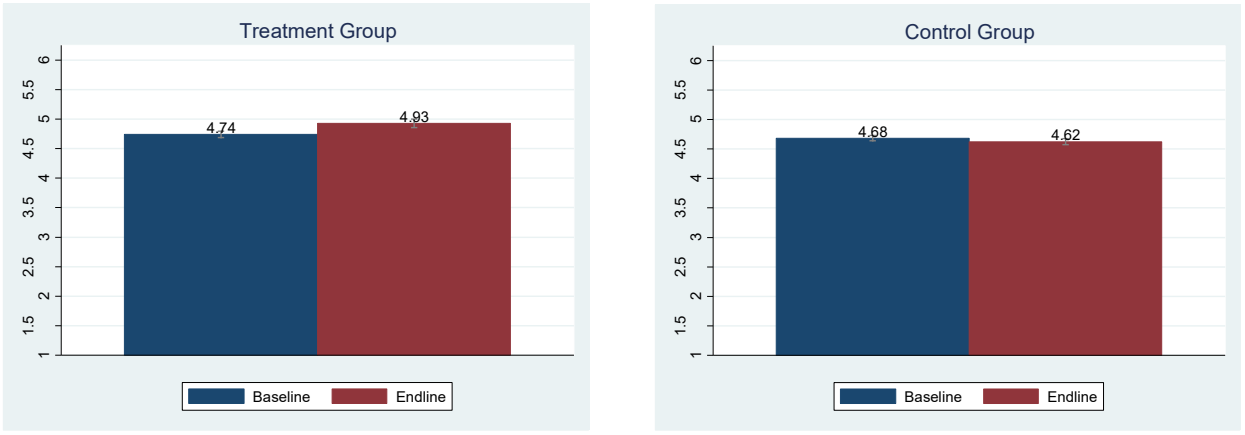

Figure 2: Baseline & Endline Mean of Final Score in Monetary Treatment and Control Groups.

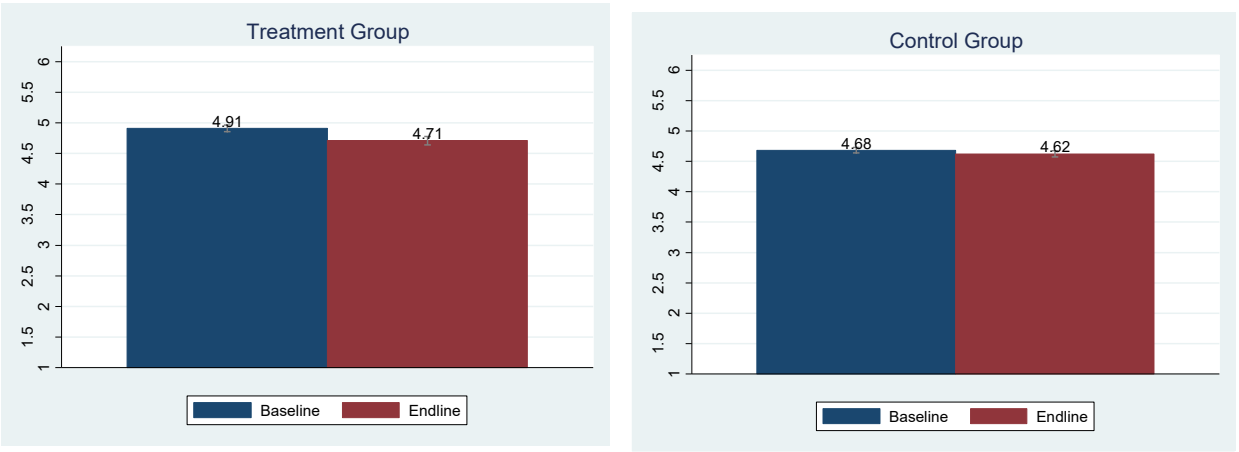

Figure 3: Baseline & Endline Mean of Final Score in Near Monetary Treatment and Control Groups.

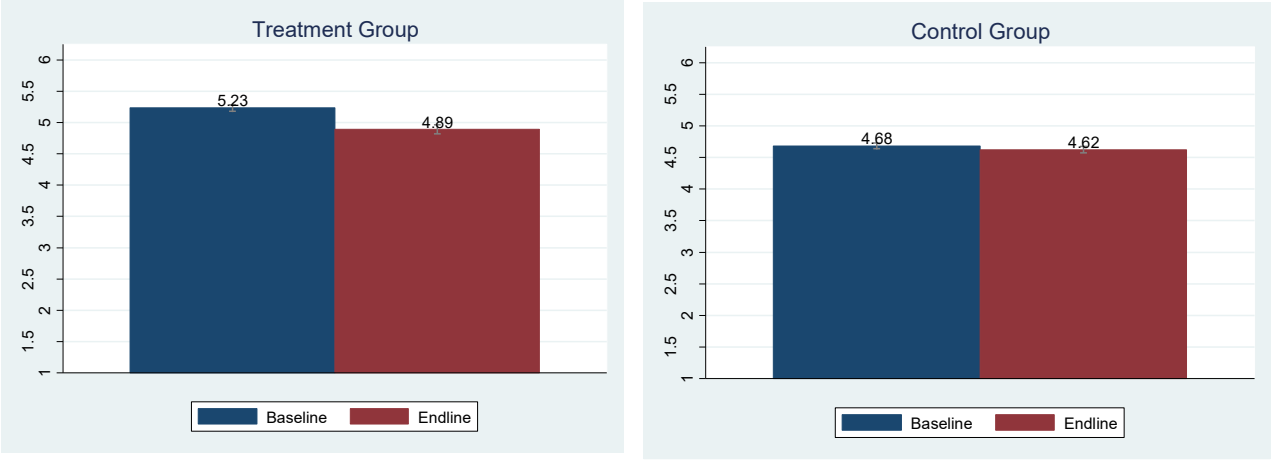

Figure 4: Baseline & Endline Mean of Final Score in Non-Monetary Treatment and Control Groups

School	District	Baseline	Endline
Aje Junior	Lagos Mainland	456	456
Falomo Junior	Eti-Osa	426	198
Falomo Senior	Eti-Osa	317	267
Iponri Junior	Lagos Mainland	328	183
Iponri Senior	Lagos Mainland	275	178
Ireti Junior	Eti-Osa	567	339
Ireti Senior	Eti-Osa	183	278
MBA Junior	Lagos Mainland	112	138
Wahab Junior	Eti-Osa	521	399
Wahab Senior	Eti-Osa	198	204
Total		3383	2640

TABLE 1: Overall Differences in Differences Estimates

	(1) Presentation of Content	(2) Clarity of Expectations	(3) Helpfulness	(4) Useful Feedback	(5) Encouraging Participation	(6) Motivation	(7) Overall Teaching Effectiveness	(8) Final Evaluation Score
Time	-0.107*	-0.074	-0.113	-0.041	0.006	-0.029	-0.164	-0.058
	(0.13)	(0.11)	(0.06)	(0.06)	(0.03)	(0.04)	(0.16)	(0.06)
Treatment	0.256	0.175	0.261	0.307	0.312	0.280	0.292	0.275
	(0.20)	(0.10)	(0.21)	(0.21)	(0.18)	(0.17)	(0.19)	(0.17)
DID	0.101	-0.160	0.077	-0.005	-0.140	-0.086	-0.084	-0.045
	(0.23)	(0.21)	(0.22)	(0.21)	(0.20)	(0.20)	(0.19)	(0.20)
_cons	4.736***	4.731***	4.707***	4.610***	4.623***	4.726***	4.658***	4.682***
	(0.07)	(0.04)	(0.07)	(0.06)	(0.03)	(0.04)	(0.10)	(0.05)
R-Squared	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03
Observations	3382	3382	3382	3382	3382	3382	3382	3382

*** Number of observations includes total number of observations at baseline and end line

TABLE 2: Monetary Differences in Differences Estimates

	(1) Presentation of Content	(2) Clarity of Expectations	(3) Helpfulness	(4) Useful Feedback	(5) Encouraging Participation	(6) Motivation	(7) Overall Teaching Effectiveness	(8) Final Evaluation Score
Time	-0.107	-0.073	-0.113	-0.041	0.006	-0.029	-0.164	-0.058
	(0.14)	(0.12)	(0.06)	(0.06)	(0.03)	(0.04)	(0.16)	(0.06)
Moneytreat	-0.047	0.094	-0.008	0.028	0.132	0.092	0.205	0.060
	(0.18)	(0.17)	(0.20)	(0.17)	(0.21)	(0.19)	(0.13)	(0.18)
DID	0.444	0.090	0.407	0.348	0.095	0.133	0.029	0.243
	(0.34)	(0.37)	(0.39)	(0.35)	(0.43)	(0.43)	(0.27)	(0.38)
_cons	4.736***	4.731***	4.707***	4.610***	4.623***	4.726***	4.658***	4.682***
	(0.07)	(0.04)	(0.07)	(0.06)	(0.03)	(0.05)	(0.11)	(0.05)
R-Squared	0.02	0.01	0.02	0.01	0.01	0.01	0.01	0.02
Observations	1920	1920	1920	1920	1920	1920	1920	1920

*** Number of observations includes total number of observations at baseline and end line

TABLE 3: Near Monetary Differences in Differences Estimates

	(1) Presentation of Content	(2) Clarity of Expectations	(3) Helpfulness	(4) Useful Feedback	(5) Encouraging Participation	(6) Motivation	(7) Overall Teaching Effectiveness	(8) Final Evaluation Score
Time	-0.107	-0.074	-0.113	-0.041	0.006	-0.030	-0.164	-0.058
	(0.14)	(0.12)	(0.06)	(0.06)	(0.03)	(0.04)	(0.16)	(0.06)
NearMonTreat	0.263	0.117	0.222	0.266	0.241	0.214	0.180	0.226
	(0.26)	(0.14)	(0.29)	(0.30)	(0.25)	(0.23)	(0.33)	(0.25)
DID	-0.096	-0.222	-0.067	-0.157	-0.165	-0.096	-0.097	-0.141
	(0.22)	(0.12)	(0.18)	(0.19)	(0.09)	(0.12)	(0.20)	(0.14)
_cons	4.736***	4.731***	4.707***	4.610***	4.623***	4.726***	4.658***	4.682***
	(0.07)	(0.04)	(0.07)	(0.06)	(0.03)	(0.05)	(0.11)	(0.05)
R-Squared	0.02	0.01	0.02	0.01	0.01	0.01	0.01	0.02
Observations	1913	1913	1913	1913	1913	1913	1913	1913

*** Number of observations includes total number of observations at baseline and end line

TABLE 4: Non-Monetary Differences in Differences Estimates

	(1) Presentation of Content	(2) Clarity of Expectations	(3) Helpfulness	(4) Useful Feedback	(5) Encouraging Participation	(6) Motivation	(7) Overall Teaching Effectiveness	(8) Final Evaluation Score
Time	-0.107	-0.073	-0.113	-0.041	0.006	-0.038	-0.164	-0.058
	(0.14)	(0.12)	(0.06)	(0.06)	(0.03)	(0.04)	(0.16)	(0.06)
NonMonTreat	0.565*	0.320**	0.582*	0.642*	0.572*	0.545*	0.495	0.549*
	(0.21)	(0.07)	(0.20)	(0.22)	(0.21)	(0.18)	(0.29)	(0.18)
DID	-0.103	-0.386	-0.166	-0.265**	-0.383***	-0.324**	-0.219	-0.283*
	(0.14)	(0.27)	(0.09)	(0.06)	(0.03)	(0.07)	(0.17)	(0.08)
_cons	4.736***	4.731***	4.707***	4.610***	4.623***	4.726***	4.658***	4.682***
	(0.07)	(0.04)	(0.07)	(0.06)	(0.03)	(0.05)	(0.11)	(0.05)
R-Squared	0.07	0.03	0.07	0.06	0.05	0.05	0.03	0.07
Observations	1889	1889	1889	1889	1889	1889	1889	1889

*** Number of observations includes total number of observations at baseline and end line