

# Teacher Opinions on Performance Incentives

Evidence from the Kyrgyz Republic

*Marlaine E. Lockheed*

The World Bank  
Europe and Central Asia Region  
Education Unit  
January 2014



## Abstract

This paper uses data from a post-hoc evaluation of a performance-based teacher incentive program in the Kyrgyz Republic to examine the opinions of teachers receiving different pay bonuses based on their performance as assessed by external evaluators. Overall, teacher opinions of the program were favorable, although teachers who received lower performance ratings held

less favorable opinions about the motivational aspects of the program. Despite this, lower-rated teachers were more likely to report that they used what they learned to evaluate their own teaching, as compared with more highly rated teachers, and were more likely to take professional development courses in the years following the program's implementation.

---

This paper is a product of the Education Unit, Europe and Central Asia Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The author may be contacted at [mlockheed@verizon.net](mailto:mlockheed@verizon.net).

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# **Teacher Opinions on Performance Incentives: Evidence from the Kyrgyz Republic<sup>1</sup>**

Marlaine E. Lockheed<sup>2\*</sup>

**JEL classification:** I28

**Keywords:** Education Policy, Teacher Incentives

**Sector Board:** Education

---

<sup>1</sup> I am grateful to the Kyrgyz Republic Ministry of Education and Sciences for authorizing and facilitating the survey, to M-Vector Consulting of Bishkek for conducting the interviews and preparing the data sets, and to the schools and teachers who participated in the surveys. In addition I wish to thank the World Bank staff who also supported this study, in particular Dingyong Hou, Paul Cahu, Aisulu Bedelbayeva, and Aisulu Sulaimonova. This paper is based on an independent evaluation supported by the Europe and Central Asia Region of the World Bank. Its findings and conclusions are the author's alone, and do not represent the World Bank Group.

<sup>2</sup> Correspondence: Marlaine Lockheed, The World Bank (retired), 561 Lake Drive, Princeton, NJ 08540, email: mlockheed@verizon.net.

## **Teacher opinions of performance incentives: Evidence from the Kyrgyz Republic**

Marlaine Lockheed

Teacher incentive programs are intended to motivate teachers to undertake behaviors that improve student learning and then to reward the teachers when the behaviors are observed or when student learning improves (Glewwe et al. 2009; Murnane & Cohen 1986; Rogers & Vegas 2009; Umansky 2005). This paper examines a teacher incentive program in the Kyrgyz Republic and asks whether teachers who received different levels of rewards held different views of the program. The remainder of the paper is organized as follows. The following section briefly reviews different types of teacher incentive programs and summarizes the few rigorous studies of their impact on student learning and teacher attitudes in developing countries. The next two sections describe the Kyrgyz Republic and the teacher incentive program that was piloted there. The fourth section outlines the data and analytic methods, and the fifth section presents the results of this analysis; the final section discusses the implications of the results.

### 1. Types of Incentive Programs

Murnane (2008) identifies three common types of incentive program for teachers: incentives to work in difficult conditions (“hardship pay”), incentives tied to mastery of knowledge or skills (“certification pay”), and incentives based on increases in student test scores (“merit pay”). Goldhaber and his colleagues (2010) add to this list incentives for teachers of selected subjects, such as science and math (“scarce-expertise pay”). Another type of incentive program studied in developing countries involves incentives for teachers to do the bare minimum: show up at school (“attendance pay”). Taken together, merit-pay, certification-pay and attendance-pay can all be considered examples of “performance

incentives” for existing teachers, whereas hardship pay and scarce-expertise pay are designed to recruit teachers for specific locations or fields of teaching. Incentive programs can be targeted at individual teachers (teacher incentive programs) or an entire school (“whole school” incentive programs); they may or may not be competitive. Competitive incentive programs reward only the “best” teachers or schools, whereas non-competitive programs reward all teachers or schools that meet pre-established standards or criteria; this paper refers to these types of programs as “standards-based” incentive programs.

### 1.1 Prior Research on Incentive Programs

No study of incentive programs in developing countries has, to my knowledge, involved either hardship pay or scarce-expertise pay, and most have involved merit pay incentives (Behrman et al. 2012; Glewwe, Holla & Kremer 2009; Glewwe, Ilias & Kremer, 2010; McEwan & Santibanez 2005; Muralidharan & Sundararaman 2011; Rau & Contreras 2009; Santiago, Cabrol & Afonso 2010). One study examined the impact of attendance pay (Duflo, Hanna & Ryan 2008), but no others to my knowledge have examined the impact of incentives conditioned on other changed teaching behaviors, which is most likely due to the difficulties of directly measuring and monitoring teaching performance (Neal 2011). Some studies, however, have looked at how “merit pay” incentives have affected teaching practices (Glewwe, Holla & Kremer 2009; Glewwe, Ilias & Kremer 2010; Lavy 2004; Muralidharan & Sundararaman 2011). Two studies examined incentive programs that rewarded teachers who received certification based on a mixture of performance indicators (Martins 2010; McEwan & Santibanez 2005). About half the incentive programs that have been studied in developing countries have involved competition. Only three prior studies have examined teacher attitudes towards any type of teacher incentive programs (Gastanadui 2013; Martins 2010, Muralidharan & Sundararaman 2011). Results from the very few teacher incentive programs that have been rigorously evaluated in developing countries have been mixed and are reviewed in the next section.

With respect to teacher attitudes, the literature, mostly from developed countries, suggests that teacher opinions of incentive programs will be more positive for standards-based incentives than for

competitive incentives, since competition necessarily produces “winners” and “losers”. Teachers typically view other teachers in their own schools as colleagues rather than as competitors, and research on “merit-based” teacher incentives has repeatedly found that teachers prefer to distribute awards equitably rather than allow for winners (Murnane & Cohen 1986). Standards-based performance assessments--despite recognizing different levels of teacher performance or “proficiency” and allowing teachers who are assessed as more “proficient” to receive larger awards than those who are assessed as less “proficient”--provide the possibility that, in principle, all teachers could be assessed as “highly proficient.” Teachers may feel that a standards-based assessment of performance is fairer than a competition.

In addition, research from developed countries suggests that teacher attitudes toward incentive programs will be unfavorable for programs that utilize “extrinsic” rewards, such as salary bonuses, and favorable for those that provide performance feedback (Deci & Ryan 2000).

## 1.2 Prior Research on the Impact of Standards-based Incentives

Individual, standards-based teacher incentives appear to be more effective than competitive teacher incentives. The only evaluated standards-based teacher incentive program designed to reward specific teaching behaviors in a developing country is a well-known randomized control trial (RCT) study of incentives designed to lower teacher absenteeism—“attendance pay” (Duflo, Hanna & Ryan 2008). Contract teachers in schools operated by an NGO in India were paid a bonus contingent upon the teacher showing up more than ten days per month; evidence of the teacher’s attendance was documented by a camera. The incentive was successful, in that the promise of a bonus (which was paid monthly) boosted teacher attendance.

Another rigorously evaluated standards-based teacher incentive program—a “merit pay” program designed to reward teachers for raising student achievement-- was carried out in India (Muralidharan & Sundararaman 2011). In this case, schools were assigned to three incentive conditions (control, individual teacher incentive and whole school (“group”) incentive) conditioned on growth in student learning, with a merit-pay bonus given to teachers for each additional percent of mean test score

improvement shown by students in the school or teacher's class, depending on the incentive condition. After five years of the incentive program, the schools in both incentive conditions registered higher growth in learning as compared with the schools in the control condition, and individual incentives to teachers were found to be more effective than the whole school incentive (Muralidharan 2012). This program, however, also led teachers to engage in after-school tutoring and additional test preparation for students (Muralidharan & Sundararaman 2011).

A third standards-based merit-pay incentive scheme was evaluated in Mexico (Behrman et al. 2012). In a rigorously designed experiment, schools were randomized into three treatment groups and a control. One of the treatments provided individual incentives to teachers on the basis of raising their student mathematics scores to specified levels. In each of three years of the program, no differences were found in student test scores between the individual teacher incentive program and the control.

### 1.3. Prior Research on the Impact of Competitive Incentive Programs

For reasons of budget constraints, competitive teacher incentive programs are often preferred, since pay bonuses can be given to a limited number of teachers up to the amount available. Again, results are mixed. In Israel, one rigorous evaluation of a within-school competition among teachers of mathematics and English found that the “merit-pay” program led to significant increases in student achievement in these subjects as well as in other subjects, through a “spill-over” effect (Lavy 2004). On the other hand, in Chile, early results from a prospective evaluation of three types of individual competitive “merit pay” programs – rewards for students on the basis of improved test scores, rewards for teachers on the basis of their students' learning gains and rewards for the principal for gains at the school level—found no impact for the competitive teacher incentive alone (Santiago, Cabrol & Afonso 2010).

In Portugal, a “certification-pay” incentive program that offered teachers a pay scale increment and promotion, based on meeting certain performance criteria, was transformed into a within-school competition due to quantity control constraints (Martins 2010). Limitations on the number of higher-level teaching positions available within each school meant that fewer teachers than all who qualified in each

school were promoted. This program was associated with a decline in student test scores. A similar program was implemented in Mexico (McEwan & Santibanez 2005). Teachers volunteered for a year-long performance assessment process, with those receiving 70 of 100 performance points becoming “certificated” for a promotion; points were based on a combination of credentials, experience, professional development, peer evaluations, subject matter knowledge and student achievement. Actual promotions, however, were competitive within schools and fewer than two-thirds of those who were eligible received a promotion. This program had no effect on improving student test scores.

The results from “merit-pay” incentive programs involving between-school competitions are also mixed. In Kenya, a two-year program provided teachers in top-ranked and most-improved schools with a salary bonus based on aggregate student performance (Glewwe, Ilias & Kremer 2003). While initially effective in raising students’ test scores, the program ultimately did not show sustained improvement in these scores (Glewwe et al. 2003). However, in Chile a study of whole-school “merit-pay” incentives combined with a between-school competition, based on school-level improvements on SIMCE (Chile’s Sistema de Medición de la Calidad de la Educación (Education Quality Measurement System)), found that the incentive program had a positive effect on growth in student test scores (Rau & Contreras 2011).

Another line of research has examined how competitive “merit-pay” incentives conditioned on improved student achievement affect teaching performance. In Kenya, whole-school competitions had no impact on improved teacher attendance or teaching practices (Glewwe, Ilias & Kremer 2004), while in Israel competitive teacher incentives increased teachers’ use of small group instruction, individual instruction and ability tracking (Lavy 2004). In Kenya and Israel, “merit-pay” incentives conditioned on student achievement led teachers to engage in after-school tutoring and additional test preparation for their students (Glewwe, Holla & Kremer 2009, Lavy 2004).

#### 1.4 Prior Research on Teacher Opinions of Incentive Programs

In developed countries, very few teachers are supportive of competitive “merit pay” schemes. For example, a study of over 3,000 teachers in the US found that fewer than 20% of teachers surveyed held

positive attitudes toward “merit pay,” and that the attitudes towards “merit pay” of teachers who received a pay incentive were no different from the attitudes of teachers who had not received a pay incentive (Goldhaber, DeArmond & DeBurgomaster 2010). By comparison, a much higher share (47%) of teachers in the same study were supportive of incentive programs based on teachers’ demonstrated mastery of content and pedagogy (“certification pay”). Research from developed countries also suggests that providing extrinsic rewards, such as money or other tangible goods, reduces intrinsic motivation, whereas positive feedback for observed competence increases intrinsic motivation (Deci & Ryan 2000).

Only three prior studies in developing countries have actually asked teachers to reflect on how a teacher incentive program affected their motivation and behavior, and the results of these studies suggest that competitive incentives produce less favorable attitudes than standards-based incentives. A study from Portugal found that teacher incentives based on within-school competitions were associated with lower job satisfaction, as indicated by early retirements (Martins 2010). By comparison, a study in India found that teachers participating in a standards-based incentive program held favorable opinions regarding standards-based salary bonuses, and the strength of that support was related to student learning gains (Muralidharan & Sundararaman 2011). And a study in Peru found that teachers held negative opinions regarding how a new teacher career law, including merit-pay, affected their motivation (Castanadui 2013).

The present study is the first in a developing country to look at the perceptions of teachers who received different pay bonus amounts from a within-school standards-based “certification” incentive program. This program in the Kyrgyz Republic awarded smaller salary bonuses to teachers who received lower performance ratings and larger salary bonuses to teachers who received higher performance ratings. In addition, teachers received performance feedback during the “certification” process, and this may have had different effects on the teachers’ motivation, depending on the rewards and feedback the teachers receive. This paper sets out to answer two questions:

(a) What are teachers' perceptions of the motivational aspects of the teacher incentive program, particularly with respect to changed teaching practices and participation in professional development?

(b) Do teachers who receive lower performance ratings and therefore lower extrinsic rewards hold different perceptions from those who receive higher ratings and higher extrinsic rewards?

This paper adds to the literature in two ways, by examining an incentive program for individual teachers that rewards their mastery of knowledge and skills ("certification pay"), and by providing new evidence about teachers' attitudes towards extrinsic incentive programs.

We use data from a post-hoc evaluation of a teacher incentive program in the Kyrgyz Republic to examine teacher awardees' perceptions of the program's motivational aspects and their self-reports of the impact of the program on their teaching behaviors. We also ask whether differences in teachers' standards-based performance ratings are associated with differences in these opinions and self-reports.

## 2. The Kyrgyz Republic

The Kyrgyz Republic is a low-income mountainous country in central Asia, with a GDP per capita of \$2,400 in 2011 and a population of 5.6 million; over a third of the population lives on income below the official poverty line (World Bank 2013). The Kyrgyz Republic's nine provinces are each divided into up to 20 districts. While literacy is widespread and primary and secondary net enrollment rates are relatively high, 88.4 and 80.4 respectively in 2011, learning achievement is poor. The Kyrgyz Republic participated in the Program for International Student Assessment (PISA) in both 2006 and 2009; its reading, mathematics and science scores for 15-year-olds were the lowest among participating countries in both years, although some improvement between 2006 and 2009 was observed (OECD 2007, 2010a, 2010b).

As in many countries, teachers in the Kyrgyz Republic traditionally were provided few incentives for good performance, once they entered the teaching profession. Uniform salary scales were largely

determined by a teacher's educational attainment and length of service, according to a formal schedule, and were quite low, averaging around about US\$42 per month in 2006 (Finland Consulting Group 2009). In the mid-2000s, the Kyrgyz Republic initiated a number of interventions designed to improve the quality of its education system, one of which was a teacher incentive program implemented in two provinces.

### 3. The Intervention and School Participation

The teacher incentive program was intended to test one way of both raising teachers' remuneration and improving their performance, by providing monthly salary bonuses to teachers depending on their level of performance as assessed through portfolio reviews, observations and interviews. Hence, it was intended as a standards-based teacher incentive program. On the basis of these standards-base performance evaluations, teachers were rated as Master Teacher (for a monthly bonus of \$17), Professional-1 (\$12.5 bonus), Professional-2 (\$12 bonus), Specialist-1 (\$9 bonus), and Specialist-2 (\$8 bonus); these bonuses amounted to salary increments ranging from 19% to 40%.

Resource limitations, however, introduced competition to the program, as within-school decisions were made to adjust the bonus payments to each school's actual resource envelope. Performance evaluations were carried out during the school year 2007/2008, and individual incentive payments were made directly to teachers' bank accounts during the school year 2008/2009. The performance evaluation was part of a broad performance management reform that included a performance management policy and performance criteria for teachers, improved guidelines and criteria for evaluating teacher performance, and a performance incentive scheme; it also involved training of school personnel in the new performance management system (Finland Consulting Group May 2009).

Schools volunteered to participate and were chosen on the basis of being considered to be "better" schools, as suggested by their use of formative assessments, the quality of the school team, the extent to which teachers used their planning diaries, and other unspecified criteria. Within schools, teacher participation was voluntary; about half the teachers in each school participated in the pilot.

The incentive program was implemented in 157 schools in two provinces—in 104 of the 199 schools in Issyk-kul province and in 53 of the 117 schools in Talas province. Participating schools included 11 percent of the nine “basic” schools offering grades 1-9 in Issyk-kul and 50% of the six basic schools in Talas, 57% of the 180 “general secondary” schools offering grades 1-11 in Issyk-kul and 46% of the 103 general secondary schools in Talas, and 12% of the four “lycees” in Issyk-kul and 75% of the four “lycees” in Talas. The average school enrolled 542 students in 2007/2008. Over three-quarters of the schools were located in communities having fewer than 5,000 residents.

#### 4. Methods and Sample

Data analyzed in this paper were collected in October 2011 by local interviewers who visited every school that had participated in the incentive program in 2007/2008. Interviewers conducted face-to-face interviews with the school principal or his/her representative<sup>1</sup> and with a random sample of teacher-awardees, stratified according to the teacher’s performance level as assessed in 2007/2008.

##### 4.1 Teacher Sample

The researchers were provided a list of teachers and instructed to select at random, in each school, one teacher to interview from each of five teacher categories; interviewers were not provided an explanation of the categories. In fact, the categories represented the five performance categories used in awarding teacher pay bonuses based on their evaluation in 2007/08: Master Teacher, Professional-1, Professional-2, Specialist-1, and Specialist-2. The intention was to interview 104 teachers in each of the five categories in Issyk-kul, for a total of 520 teachers, and 53 teachers in each of the categories in Talas, for a total of 265 teachers, but this intention was not realized, due to the actual distribution of awardees across categories (table 1). Because very few teachers received either the highest Master Teacher award or the lowest Specialist-2 award, teachers in these categories were missing in many schools. However, in both provinces, the researchers were able to interview over 30 percent of teachers in most categories, reaching 505 teachers in Issyk-kul (28 percent of program teacher population) and 261 teachers in Talas

(34 percent of the program teacher population). Interviews with “others” not on the list of awardees comprised 2.8 percent of the interviews in Issyk-kul and 3.8 percent of the interviews in Talas.

The typical teacher-awardee was a middle-aged, Kyrgyz-speaking woman with a university degree in education and over 20 years experience teaching; about half had a “higher qualification” level at the outset of the incentive program in 2007/2008 (table 2). Teacher-awardees in Issyk-kul were slightly more likely to hold a higher teaching rank and were slightly less likely to teach elementary subjects than those in Talas, but in other respects the teachers in the two provinces were quite similar. Compared with the national samples of grade 4 and grade 8 teachers surveyed as part of the National Sample-Based Assessment (NSBA), the present sample was less likely to hold a “higher teaching qualification” but was comparable to the NSBA in terms of percent female, education and experience (CEATM 2010).

#### 4.2 Questionnaires and Interview Forms

School and teacher questionnaires were prepared in English and Russian and then translated into Kyrgyz (from Russian, with the English versions providing additional guidance for translation). The school interview included 18 multi-part questions, took about 15 minutes to conduct, and asked the principal about the structure and organization of the school, the students and teachers, special programs the school had participated in, and school policies and practices. Six questions addressed the teacher incentive program specifically. The teacher interviews included 40 multi-part questions (some of which were open-ended), took 20-30 minutes to conduct, and asked teachers about: (a) the usefulness and effectiveness of the Teacher Incentive/Performance Management program, (b) the teacher’s participation in the program, and (c) the teacher’s background.

#### 5. Results

The incentive program was intended to motivate teachers, build skills and ultimately improve teacher classroom performance, for which teachers would be rewarded. This section presents evidence regarding the opinions and reports from teachers who participated in the program in 2007/2008 and who

received a salary bonus as a result of this participation; these teacher-awardees are referred to as teachers in the rest of this paper.

### 5.1 Teacher Motivation

Overall, 91 percent of teachers thought the purpose of the incentive program was to increase professional motivation—that is, to encourage teachers to work harder; only 7.5 percent thought its purpose was to reward good teachers; thus, they did not view the teacher incentive program as a “merit pay” scheme. Teachers reported high levels of motivation from the program. For example, in response to the question “How would you rate the effectiveness of the pilot implementation of the teacher incentive/performance management program in increasing your motivation to take professional development courses”, 86 percent of the teachers answered “effective” or “extremely effective.”

From the teacher survey, we created a *teacher motivation index*, which was the sum of the ratings (from “not effective” to “extremely effective”) given by teachers to questions regarding their opinions regarding the extent to which the program increased their motivation to change their teaching practices, to take professional development courses and to continue working as a teacher. On average, there were no statistically significant differences among teachers receiving different performance ratings with respect to their opinions on the motivational aspects of the incentive program, with scores averaging around 9.2 points on a 12-point scale for all categories of teacher: Master Teacher (9.32), Professional-1 (9.30), Professional-2 (9.24), Specialist-1 (9.24) and Specialist-2 (9.19).

We next used OLS regression analysis to adjust for individual and location differences among respondents; the results are presented in columns 1 and 2 of table 3, once for performance rating as a scale and once for each category of performance rating. After controlling for district school location, teacher’s age, gender, experience and high qualification rank in 2007, we found that teachers who had received lower ratings in 2007 (Specialists 1 and 2) held less positive opinions in 2011 regarding the motivational aspects of the program, as compared with Master Teachers.

Finally, we looked at each of the questions that comprised the motivation index separately, to explore this heterogeneity. We found that teachers who received higher ratings reported being more motivated to change their teaching practices ( $X^2_{(6)} = 17.71, p = .023$ ), but did not recall being more motivated to take professional development courses or continue working as a teacher, in comparison with teachers who received lower ratings.

## 5.2 Changed Teaching Practice

But did these more highly rated teachers report that they actually changed their teaching practice? In general, teachers reported that they used the skills imparted through the program. For example, in response to the question “How much have you used the knowledge, skills or information that you acquired from the pilot implementation of the teacher incentive/ performance management program in evaluating your own teaching practices”, 76 percent of the teachers answered “often” or “very much.” We also created an index of *teacher use of knowledge and skills*, which was the sum of the ratings (from “little or not at all” to “very much”) given by teachers to questions regarding the frequency of their use of the knowledge or skills acquired through the program related to monitoring the learning progress of students, doing formative assessments, changing their own teaching practices and evaluating their own teaching knowledge and skills. On average, there were no differences among teachers receiving different performance ratings with respect to how much they used the knowledge and skills from the program. The average index score was about 16 points on a 20-point scale for all performance groups: Master Teachers (16.12), Profession-1 (16.33), Professional-2 (16.30), Specialist-1 (16.03) and Specialist-2 (16.47). Importantly, the Master Teachers who received the highest ratings, and hence the highest bonuses, did not report using the knowledge and skills more than those who received a lower rating and monetary award.

Since older, more experienced teachers might not be as responsive to training in new instructional practices and the average age of teachers was higher for each higher performance group<sup>2</sup>, we adjusted for these individual characteristics, as well as for school location differences, through OLS regressions; the results are presented in columns 3 and 4 of table 3, once for performance rating as a scale and once for

each category of rating. After controlling for district school location, teacher's age, gender, experience and higher qualification level in 2007, we still found no relationship between a teacher's performance rating in 2007 and his or her reported changes in teaching behaviors.

Finally, we looked at each of the components of the *teacher use* index to see if teachers' performance ratings were associated with higher use in any particular area. We found that teachers who received higher ratings did not report greater frequency in using the skills encouraged by the program (monitoring the learning progress of their students and using formative assessments), or report more frequent use of the knowledge, skills and information from the program to change their teaching practices, in comparison with teachers receiving lower ratings. However, teachers who received lower ratings reported more self-evaluation of their own teaching knowledge and skills ( $X^2_{(8)} = 15.81, p = .045$ ).

### 5.3 Professional Development

We then investigated whether teachers who received lower ratings were more likely to have taken professional development courses after they received their ratings, in an effort to improve their knowledge and skills. As shown in table 4, column 1, teachers who received lower performance ratings were more likely to have taken professional development courses in the years following the program's implementation. Most of this effect was for teachers in the lowest performance rating category, who were much more likely to have taken recent professional development courses, as compared with master teachers (table 4, column 4); 91 percent of Specialist-2 teachers reported having taken recent professional development courses, as compared with fewer than 75 percent of teachers in all other performance categories ( $X^2_{(4)} = 13.26, p = .01$ ).

## 6. Summary and Discussion

Overall, teachers recalled that the teacher incentive program was motivating and provided useful skills and knowledge. A few differences in perceptions regarding the program were found among teacher

awardees, by level of award. Teachers who received lower performance ratings held less favorable opinions about the motivational aspects of the incentive program. Despite this, lower-rated teachers were more likely to report using what they learned to evaluate their own teaching, as compared with more highly rated teachers, and were more likely to take professional development courses in the years following the program's implementation.

In other respects, teachers' performance ratings were uncorrelated with their: (a) recall of using new teaching methods, such as monitoring student learning and using formative assessments, (b) report of having more positive attitudes regarding their career prospects, or (c) report of being motivated to continue teaching. These results are generally consistent with findings from recent studies of teacher incentives in the United States (Yuan et al. 2013, Springer et al. 2012).

Since the Kyrgyz Republic's teacher incentive program is best described as a standards-based performance incentive program rather than as a competitive incentive program, the overall results are also consistent with the previous studies in India and the US, with respect to overall favorability of teachers' opinions. However, the somewhat more negative attitudes of lower rated teachers – with respect to the motivational aspects of the program -- could reflect the within-school competitive aspect of the process. The overall budget for incentives in the program was insufficient for all teachers to receive Master Teacher salary bonuses, even if they had qualified. Although lower-rated teachers received a bonus, their bonus was not as large as that received by the higher-rated teachers, which may have affected their favorability ratings.

Other things may have accounted for the lack of heterogeneity in program effects, as reported by teachers. First, the teachers were asked to reflect on the program approximately four years after their participation, which may have diminished their recall and eliminated the likelihood of finding differences among teacher-awardees. Second, the teacher incentive program was followed by a "whole school" incentive program, whose effects may have diluted teachers' recall about the individual teacher incentive program that had preceded it. Third, the salary bonuses were relatively large for all groups, which could account for the overall positive opinions for all teacher awardees, even those whose bonuses were on the

lower end. A comparison group of teachers who did not participate in the performance evaluation or who participated but did not receive an award might have revealed less positive opinions and would have been preferable, but the identity of such teachers was not made available to the research team.

In general, however, the teachers who received awards viewed the standards-based incentive program favorably, and their favorability ratings were significantly higher than those given to competitive teacher incentive programs in other countries.

The present study has three implications for future teacher incentive programs. First, teacher incentive programs should be standards-based rather than competitive. Second, teacher incentive programs should avoid a “two-pronged” approach that establishes qualification standards but then awards bonuses competitively, since this may affect both teacher satisfaction and the program’s impact on student performance. Third, teacher incentive programs that award merit pay should establish bonus levels that are affordable, even if all teachers attain the performance standard.

#### Notes

1. Half the respondents were school principals, and approximately half were heads of curriculum; four respondents were deputy directors.
2. The mean ages were: Master Teachers, 50.8 years; Professional-1, 49.1 years, Professional-2, 45.3 years; Specialist-1, 39.1 years; Specialist-2, 36.6 years.

## REFERENCES

Behrman, J., S. Parker, P. Todd, K. Wolpin (2012 October). “Aligning learning incentives of students and teachers: Results from a social experiment in Mexican high schools” University of Pennsylvania, Department of Economics.

Center for Educational Assessment and Teaching Methods (CEATM) (2010). *Report on the National Evaluation of Educational Achievements of Pupils, 2009*. Bishkek: CEATM.

Contreras, D. and T. Rau (2012). Tournament incentives for teachers : Evidence from a scaled-up intervention in Chile. *Economic Development and Cultural Change* 61(1): 219-246.

Cantrell, S., J. Fullerton, T. Kane and D. Staiger (2008). “National Board Certification and Teacher Effectiveness: Evidence from a Random Assignment Experiment”. NBER Working Paper 14608. National Bureau of Economic Research, Cambridge, Mass.

Deci, E. and R. Ryan (2000). The “What” and “Why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11 (4): 227-268.

Duflo, E., R. Hanna and S. Ryan (2008). “Monitoring works: Getting teachers to come to school”. Center for Economic Policy Research Discussion Paper 6682.

Finland Consulting Group (draft 22 May 2008). “Rural Education Project, Component 1: Improving Teacher Incentives, Final Evaluation of Pilot Study”

Gastanadui, L. (2013). “Teacher evaluation and resistance to change: A mixed-method study of the Peruvian new teacher career law.” Unpublished doctoral dissertation, University of Texas at Austin.

Glewwe, P., A. Holla, and M. Kremer (2009). Teacher incentives in the developing world. In Matthew G. Springer (Ed.) *Performance Incentives: Their Growing Impact on American K-12 Education*. Washington, DC: Brookings Institution Press.

Glewwe, P., N. Ilias and M. Kremer (2003). "Teacher Incentives". NBER Working Paper 9671. National Bureau of Economic Research, Cambridge, Mass.

Goldhaber, D., M. DeArmond and S. DeBurgomaster (2010). "Teacher Attitudes about Compensation Reform: Implications for reform implementation. CALDER Working Paper 50. Urban Institute, National Center for Analysis of Longitudinal Data in Education Research, Washington, DC.

Lavy, V. (2004). "Performance Pay and Teachers' Effort, Productivity, and Grading Ethics." NBER Working Paper 10622. National Bureau of Economic Research, Cambridge, Mass.

Lopez-Acevedo, G. (2004). "Professional Development and Incentives for Teacher Performance in Schools in Mexico". World Bank Policy Research Paper 3236. Washington, DC: The World Bank.

Martins, P. (2010) "Individual Teacher Incentives, Student Achievement and Grade Inflation" . London School of Economics, Centre for the Economics of Education.

McEwan, P. and L Santibanez (2005). Teacher and principal incentives in Mexico. In E. Vegas (Ed.), *Incentives to Improve Teaching*. Washington, DC: The World Bank.

Muralidharan, K. (2012). "Long-Term Effects of Teacher Performance Pay: Experimental Evidence from India."

Muralidharan, K. and V. Sundararaman (2011). Teacher Opinions on Performance Pay: Evidence from India. *Economics of Education Review* 30 (1): 394-403.

Muralidharan, K. and V. Sundararaman (2009). "Teacher Performance Pay: Experimental evidence from India" NBER Working Paper 15323. National Bureau of Economic Research, Cambridge, Mass.

Murnane, R. J., and D. K. Cohen (1986). Merit Pay and the Evaluation Problem: Why most merit pay plans fail and a few survive. *Harvard Education Review* 56: 1-17.

Murnane , R. J., (2008) “Educating Urban Children” NBER Working Paper 13791. National Bureau of Economic Research, Cambridge, Mass.

Neal, D. (2011). “The Design of Performance Pay in Education”. NBER Working Paper 16710. National Bureau of Economic Research, Cambridge, Mass.

OECD (2007). *PISA 2006: Science Competencies for Tomorrow’s World*. Paris: OECD

OECD (2010a). *PISA 2009: What Students Know and Can Do--Student Performance in Reading, Mathematics and Science*. Paris: OECD

OECD (2010b). *Kyrgyz Republic 2010: Lessons from PISA*, Paris: OECD

Rogers, H. and E. Vegas (2009). “No more cutting class: Reducing teacher absence and providing incentives for performance.” Policy Research Working Paper 4847. Washington, DC: The World Bank

Santiago, A., M. Cabrol and M. Alfonso (February 2010). Aligning Learning Incentives (ALI). Inter-American Development Bank, Education Division. <http://www.iadb.org/en/topics/education/aligning-learning-incentives-ali,7466.html>

Springer, M., J. Pane, V. Le, D. McCaffrey, S. Burnes, L. Hamilton and B. Stecher (2012). Team Pay for Performance: Experimental evidence from the Round Rock pilot project on team incentives. *Educational Evaluation and Policy Analysis* 34(4): 367-390.

Umansky, I. (2005). A Literature Review of Teacher Quality and Incentives: Theory and evidence. In E. Vegas (Ed.), *Incentives to Improve Teaching*. Washington, DC: The World Bank.

World Bank (2004). Kyrgyz Republic Rural Education Project: Project Appraisal Document.

World Bank (2013) data on Kyrgyz Republic

Yuan, K., V. Le, D. McCaffrey, J. Marsh, L. Hamilton and B. Stecher , Springer, M., (2012). Incentive Pay Programs Do Not Affect Teacher Motivation or Reported Practices: Results from Three Randomized Studies. *Educational Evaluation and Policy Analysis* 35(1): 3-22.

**Table 1. Population of teacher incentive awardees and share of teacher-awardees interviewed, by teacher performance level, 2011**

Performance level	Issyk-kul			Talas		
	Population	Achieved sample	As a share of population	Population	Achieved sample	As a share of population
Master Teacher	177	61	0.34	16	8	0.50
Professional-1	898	225	0.25	236	76	0.32
Professional-2	445	121	0.27	274	83	0.30
Specialist-1	168	51	0.30	158	55	0.35
Specialist-2	93	29	0.31	92	29	0.32
Other	0	18	--	0	10	--
Total	1781	505	0.28	776	261	0.34

Source: Socium Consult, World Bank/M-Vector Teacher Survey 2011

**Table 2. Characteristics of incentive program teacher-awardees interviewed, Issyk-kul and Talas, 2011, compared with national sample, 2009**

Teacher Characteristics	Teacher Survey 2011			NSBA 2009	
	Total	Issyk-kul	Talas	Grade 4	Grade 8
Percent female	93	93	93	98.1	89.0
Average age in years	45.7	46	45	--	--
Percent speaking Kyrgyz at home	91	89	93	--	--
Education—Percent with university-level specialty in pedagogy	91	92.5	88.5	81.3	92.3
Experience--Percent with more than 15 years teaching experience	75.2	78	69.7	81.3	67.0
Teaching rank in 2007/2008-- Percent with “higher qualification”	54.0	58.0	46.7	73.8	65.3
Percent elementary school teacher	26.9	24.0	32.6	88.8	4.0

Source: Socium Consult, World Bank/M-Vector Teacher Survey 2011

**Table 3. OLS regression results for determinants of teacher perceived motivation and reported use of knowledge and skills**

Independent variables	Motivation		Use of knowledge, skills	
	(1)	(2)	(3)	(4)
Performance rating (scale)	-.141** (.068)		-.017 (.107)	
Master Teacher (omitted)		-----		-----
Professional 1		-.225 (.216)		-.035 (.336)
Professional 2		-.406 (.233)		-.100 (.365)
Specialist 1		-.498* (.279)		-.274 (.435)
Specialist 2		-.560* (.319)		.125 (.500)
Age in years	.002 (.011)	.003 (.011)	.019 (.017)	.020 (.017)
Male	-.382 (.239)	-.391 (.240)	-.219 (.373)	-.216 (.374)
Experience > 25 years	-.126 (.174)	-.130 (.175)	-.234 (.270)	-.259 (.272)
Higher qualification in 2007	-.023 (.158)	-.023 (.158)	.156 (.244)	.138 (.245)
Constant	9.484 (.543)***	9.400 (.527)***	14.973 (.852)***	14.974 (.827)***
Number of observations	733	733	736	736
R <sup>2</sup>	.08	.09	.097	.099

\*\*\* p < .01; \*\* p < .05; \* p < .10; Source: World Bank/M-Vector Teacher Survey 2011

Regional dummy variables are included, but not reported

**Table 4. Logit regressions results for taking a professional development course, post pilot**

Independent variables	Recent Professional Development	
	(1)	(2)
Performance rating (scale)	.189*	
	(.102)	
Master Teacher (omitted)		----
Professional 1		-.468
		(.323)
Professional 2		-.446
		(.345)
Specialist 1		-.048
		(.418)
Specialist 2		1.278**
		(.591)
Age	-.012	.003
	(.016)	(.011)
Male	.531	.479
	(.388)	(.390)
Experience > 25 years	.079	-.030
	(.252)	(.202)
Higher qualification in 2007	-.199	-.240
	(.232)	(.221)
Constant	1.769	2.056
	(.813)***	(.431)***
Number of observations	738	738
Pseudo R <sup>2</sup>	.07	.09
Chi-sq	61.67***	74.74***

\*\*\* p < .01; \*\* p < .05; \* p < .10; Source: World Bank/M-Vector Teacher Survey 2011

Note: Regional dummy variables are included, but not reported