

Summary of Recent Findings on the Impact of Performance Pay Incentives

Description of the Most Rigorous Studies

While research has been conducted on performance pay for some time, only recently have there been studies using rigorous experimental approaches. Experiments that use randomized assignment of educators into treatment and control groups (treatment being the opportunity to receive an incentive) offer the best assessment of the direct effect of incentives on student and teacher performance outcomes. The best evidence on pay for performance comes from these studies:¹

- *Springer et al. (2010)*: The POINT project in Nashville was a three-year randomized experiment involving fifth to eighth grade math teachers that rewarded teachers for certain value-added scores based on student achievement. Bonuses of \$15,000, \$10,000, and \$5,000 were awarded for teachers whose value-added scores were at historical 95th, 90th, and 80th percentiles, respectively.
- *Fryer et al. (2011/2013)*: A two-year randomized experiment for high need K-12 schools in New York, as defined by poverty rates, demographic characteristics, and test scores. Bonuses were distributed on the basis of school Progress Report scores, which are calculated on the basis of student test scores and Regents Examinations, graduation rates (for high schools), student attendance, and school environment. Schools could earn up to \$3,000 per teacher for meeting targets; distribution of money was up to the discretion of the school.
- *Fryer et al. (2012)*: A one-year randomized experiment of “traditional” (money given to teachers at the end of the year for meeting targets) and “loss aversion” (money given to teachers at the beginning of year that could be kept if a teacher met targets; otherwise, the money had to be returned at the end of the year) bonuses among math and reading K-8 teachers in Chicago Heights, Illinois. Team and individual bonuses were compared. Up to \$8,000 was available for bonuses, with the average bonus payout being \$4,400.
- *Glazerman & Seifullah (2012)*: A combination experimental and quasi-experimental study of Chicago’s Teacher Advancement Program (TAP) in 34 schools (17 control, 17 experiment). Bonuses ranged between an average of \$1,100 to \$2,500 per year, with a maximum of \$2,700 to \$6,400, depending on the year.
- *Springer et al. (2012)*: Two-year experimental study of team incentives for Round Rock middle school teachers in Texas. Bonuses (approx. \$5,900) were awarded on the basis of

¹ There is research by Goldhaber and Walch (2012) that speaks to the effectiveness of another popularly-cited program—ProComp, a five-year pay for performance program in Denver. A review of ProComp is not included in this brief, however, because the research does not provide a direct answer to the question of *performance incentive effectiveness*. There are two reasons why this is the case. First, teachers were not randomly assigned to treatment or control groups, so the research cannot find a direct *Intent to Treat* effect, and second, ProComp was more than just pay for performance, including comprehensive professional development, upgrades to information systems, updated teacher evaluation systems, and changes to district personnel. It is therefore difficult to assess the impact of the performance pay component associated with ProComp from the research available because the research assesses the impact of ProComp in general.

Summary of Recent Findings on the Impact of Performance Pay Incentives December 2013

student growth (value-added), if a team's score ranked in the top third of treatment group teams in the same grade level.

Additionally, recent quasi-experimental research also has offered insight into the impact of performance incentives for teachers who are just below or just above an “effective” threshold.²

- *Dee and Wyckoff (2012)*: A regression discontinuity study that examines the impact of teacher retention and performance in Washington D.C.'s high-stakes, district-wide IMPACT program. The measurement used to assess teachers was a combination of formal observations and value-added scores, producing an overall IMPACT score ranging from 100 to 400. After the first year, teachers with an IMPACT score below 175 were let go. Teachers with a score between 175 and 249 (labeled “Minimally Effective”) were subject to dismissal the following year should they not receive an “Effective” (250-349) or “Highly Effective” (350+) score the following year. Additionally, teachers receiving a Highly Effective score could receive a one-time bonus of \$25,000, and two consecutive “Highly Effective” scores would earn teachers substantial base pay rates resulting in as much as \$27,000 additional base pay per year.

There is also an analysis of performance incentives on teacher practices based on randomized experiments:

- *Yuan et al. (2013)*: An analysis of teacher practice survey results from three randomized experiments in three states (Tennessee, New York, and Texas). Tennessee and New York studies come from POINT and SPBP projects, respectively (see Springer 2010 and Fryer 2011, above). The Texas “PPTI” program awarded team-based bonuses on the basis of value-added to middle school teacher teams consisting of at least one teacher for the core subjects of math, reading/English, science, and social studies. Average bonuses ranged from \$5,300 to \$5,800 in the Texas program.

Findings from the Most Rigorous Studies on Performance Pay

1. *There is little evidence that traditional-model performance incentives increase student outcomes.*

Student Achievement. In general, the Nashville POINT, New York, Chicago Heights, Chicago TAP, and Round Rock, Texas programs found statistically insignificant effects of the opportunity to earn a “traditional” (money given to teachers at the end of the year for reaching performance target) bonus on student test scores. However, there are a few conditions that produced significant results, albeit not always positive results. Researchers in Nashville found a slightly positive effect of incentives on fifth grade performance in years two and three of their study; researchers in New York found slightly negative impacts of incentives in middle schools math.

Student Behavior. The evidence from New York suggests that there are insignificant effects of the opportunity to earn a bonus on student attendance, behavior, or grades (Fryer 2011).

² Quasi-experimental regression discontinuity designs assume that assignment to just below versus just above a cut-off point is essentially random, which, though not as ideal as a random control trial, offers a rigorous study design.

2. *In general, it does not matter whether the bonus is individual or team/school-based; the results appear to be the same.*

POINT (Nashville) and IMPACT (Washington, D.C.) bonuses were individual, but New York and Chicago TAP bonuses in the first year were largely school-based. Round Rock and Chicago TAP bonuses in the second year were team-based, and in Chicago Heights, “loss” and traditional-style bonuses produced the same effects, whether they were individual or team-based awards.

3. *There is little evidence that traditional-model performance incentives change teacher behavior.*

According to survey results and researchers’ validity checks, teachers in Nashville, New York, and Texas (both Round Rock and PPTI) did not change their teaching practices as a result of bonuses. Specifically, there was no evidence that incentives affected retention or teacher absences, nor did teachers report changes to their instruction (e.g., teaching to the test), number of hours worked, stress, or collegiality as a result of the bonus program. Note, though, that POINT treatment teachers in Nashville did report greater emphasis on test preparation and collaboration than did their control group counterparts. Additionally, with the knowledge that two “Minimally Effective” ratings in a row would result in job loss, a “Minimally Effective” label for teachers in the first year of the Washington, D.C. IMPACT program increased attrition by 50%.

4. *It is unlikely that traditional bonuses do not work only because they are too small.*

Up to \$15,000 was available for teachers in the POINT project, with \$10,000 and \$5,000 available at lower student achievement thresholds; up to \$8,000 was available in the Chicago Heights project; and the mean bonus distributed in Texas was between \$5,300 and \$5,900.

5. *Though there is little evidence that traditional bonuses have affected teacher or student performance, there is some evidence that “loss-aversion” bonuses—or bonuses that are given at the beginning of the year and are taken away if students do not meet expected test criteria—increase student achievement.*

According to Fryer’s (2012) Chicago Heights project, students whose teachers were in the “loss” treatment group showed large and statistically significant gains between .2 and .4 standard deviations in math, which is between 6.8 and 11 percentile points. The effects of loss-aversion bonuses are the same whether these bonuses are offered as individual or team-based bonuses.

6. *Similarly, in one high-stakes case, there is some evidence that effectiveness labels—and perhaps by extension, the opportunity to receive a bonus—may be effective in improving the overall teacher workforce through voluntary attrition and increased teacher performance, particularly for teachers near the threshold of effectiveness categories.*

As mentioned briefly above, according to Dee and Wyckoff (2012), at the threshold of being labeled minimally effective versus ineffective, voluntary attrition for minimally effective teachers increased 50%. However, among those minimally effective teachers who stayed, a “Minimally Effective” rating increased teacher performance by .27 SD. For high performing teachers, financial incentives did not improve retention but did improve teacher performance. Teachers at the “Highly Effective” threshold improved teacher performance by .24 SD.

Summary of Recent Findings on the Impact of Performance Pay Incentives
December 2013

Summary of Recent Studies

Study	Location of Program	Performance Measure	Number of Schools & Teachers	Size of Bonus	Who is Evaluated	Length of Study	Type of Design	Findings	Notes
Springer et al. (2010) POINT Project	Nashville, TN	Value added (subtracted state wide average on TN test gain from student's own annual gain and averaged these scores over the class). Bonuses given for reaching 95 th , 90 th , and 80 th percentiles of historical distribution	296 teachers in year one, 148 by year three. Grades 5-8 only	Up to \$15,000, with opportunities for \$10,000 and \$5,000 at lower thresholds	Indiv. teachers	3 years	RCT	No overall effect of incentive on student performance. Some positive effect of bonus on fifth grade teachers in years two and three only. Little evidence that incentives changed teachers' professional practice.	High attrition rates. 33% of treatment group received bonuses. Average bonus ranged between \$9,000 and \$11,000. Even control group offered small \$750 stipend to participate
Fryer et al. (2011/2013) New York Incentives program	New York City	"School Report Card," composed of state test scores and regent exams, graduation rates, student attendance, and learning environment surveys administered to teachers, parents, and students	198 schools year one, 189 year two	Up to \$3,000 per staff member per school and \$1,500 if school met 75% of target. Bonuses distribution determined by school's discretion	School	2 years	RCT	Incentives do not affect student achievement (tests or grades), student behavior (attendance or behavior), or teacher behavior (retention, absences, or responses to learning survey).	Most schools opted for group incentive scheme.

Summary of Recent Findings on the Impact of Performance Pay Incentives
December 2013

Study	Location of Program	Performance Measure	Number of Schools & Teachers	Size of Bonus	Who is Evaluated	Length of Study	Type of Design	Findings	Notes
Fryer (2012) Chicago Loss Aversion	Chicago Heights, IL	Students' end of year performance, "pay by percentile" method—teachers receive average percentile of students' ranks within their bins of 9 nearest neighbors	150 K-8 teachers, 9 schools	Up to \$8,000, average of \$4,400. For half of treatment group, received possibility of up to \$8,000 at end of semester. For other half, received \$4,000 up front with potential to earn up to \$4,000 more, had to return the difference at end of year if didn't meet target	Indiv. and teams (not clear how teams are defined)	1 year	RCT	Students whose teachers were in the "loss" treatment group had large and statistically significant gains of .2 to .4 SD in math. There are no impacts of "gains" incentives, and there is little difference between individual or team treatment within loss or gain bonuses.	Chicago Heights is a primarily low-income, high minority district. Unclear whether the study's statistical power is sufficient.
Glazerman and Seifullah (2012) Chicago TAP	Chicago, IL	Outcome of interest is student Achievement. Bonuses awarded based on formal observations of teachers and school value added (first two years) and school-grade team value added (last two years)	34 Schools in Chicago	On average, between \$1,100 and \$1,900, with a maximum payout of \$2,700-\$6,400, depending on the year	Schools and Teams of Teachers	2 years in this study (4 total years)	RCT and Quasi-Experiment	No significant or consistent effects of opportunity to participate in TAP on student achievement. Some positive, though inconsistent, effects on teacher retention.	Poorly-implemented program. Awards were not as high as intended to be, and value-added data could not be linked with student rosters.

Summary of Recent Findings on the Impact of Performance Pay Incentives
December 2013

Study	Location of Program	Performance Measure	Number of Schools & Teachers	Size of Bonus	Who is Evaluated	Length of Study	Type of Design	Findings	Notes
Springer et al. (2013) Round Rock	Round Rock, TX	Achievement growth (value added, difference in expected versus actual student performance) in four core subjects (math, reading, social studies, science). Bonuses awarded to top third of treatment groups in same level	159 teams of teachers (665 total teachers) grades 6-9 in 9 middle schools	Up to \$5,500 in year one and \$6,000 in year two; prorated as low as \$3,800 in year one and \$4,200 in year two	Teams of teachers	2 years	RCT	No effect of performance incentives on student achievement in any subject area over the two years, no effect on perception of team dynamics, no effect on teacher attitudes or practices.	
Dee and Wyckoff (2012) IMPACT	Washington, D.C.	Outcome of interest: teacher retention and effectiveness, measured by value added and formal observations into an IMPACT score ranging from 100 to 400	The entire Washington, D.C. school district	\$25,000 one time bonus and up to \$27,000 annual base pay rate per year	Individual teachers	2 years	RD	At the threshold, for those labeled minimally effective in year one, there was an increase in voluntary attrition by 50%, but for those who chose to stay, a “Minimally Effective” rating increased teacher performance the following year by .27 SD. Evidence that the highest rating (“Highly Effective”) increased teacher performance by .24 SD.	Limitation due to no random assignment, but the large-scale application and high-stakes nature of program offers interesting insight

Summary of Recent Findings on the Impact of Performance Pay Incentives
December 2013

Study	Location of Program	Performance Measure	Number of Schools & Teachers	Size of Bonus	Who is Evaluated	Length of Study	Type of Design	Findings	Notes
Yuan et al. (2013) Performance Pay on Teacher Behavior	Nashville, New York City, & Texas	Outcome of interest is teacher practices, measured by surveys	About 500 teachers in Nashville and Texas, along with 427 schools in New York	Up to \$15K for Nashville, \$3000 per teacher in New York, and about \$5000 in Texas	Indivs. in Nashville, schools in New York, and teams in Texas	1 to 3 years	RCT	Most teachers do not report that their program was a motivator for performance. None of the three programs changed teachers' instruction, increased number of hours worked, reduced job stress, or damaged collegiality.	

Works Cited

- Dee, Thomas and James Wyckoff. (2013). *Incentives, selection, and teacher performance: Evidence from IMPACT*. NBER Working Paper No. 19529.
- Fryer, Roland G. (2013). Teacher incentives and student achievement: Evidence from New York City Public Schools. *Journal of Labor Economics* 31(2): 373-407.
- Fryer Jr., Roland G, Steven D, Levitt, John List, and Sally Sadoff. (2012). *Enhancing the efficacy of teacher incentives through loss aversion: A field experiment*. NBER Working Paper No. 18237.
- Glazerman, S. and A. Seifullah. (2012). *An evaluation of the Chicago Teacher Advancement Program (Chicago TAP) after four years*. Princeton: Mathematica Policy Research.
- Springer, M.G., Ballou, D., Hamilton, L., Le, V., Lockwood, J.R., McCaffrey, D., Pepper, M., and Stecher, B. (2010). *Teacher pay for performance: Experimental evidence from the Project on Incentives in Teaching*. Nashville, TN: National Center on Performance Incentives at Van-derbilt University.
- Springer, M., J. Pane, V. Le, D. McCaffrey, S. Burns, 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.
- Yuan, Kim, Vi-Nhuan Le, Daniel F. McCaffrey, Julie A. Marsh, Laura S. Hamilton, Brian M. Stecher, and Matthew G. Springer. (2013). 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.