

Evaluating Teacher Performance Incentives While Enhancing Equity

Victor Lavy

The Hebrew University of Jerusalem, CEPR and NBER

2nd European Symposium on Economics of Education

“Efficiency and Equity in European
Education and Training Systems”

November 2005

Policy Context

- Education is very important for social development and economic growth
- Key determinant of earnings of individuals and income per capita of nations
- Consensus that teachers are likely to lead to better schools and educational outcomes
- But the question of how to attract better teachers is unresolved.

Claims

- Performance in education is hampered by a failure to provide appropriate incentives
- Schools do not face market pressure as in private sector
- Failure to allow competition among schools hinders productivity and raises costs
- Teacher and school administrator incentives have been distorted as a result of the same failure to face competition in the market
- Compensation for teachers is not “results oriented”
- The lack of financial recognition of teaching performance contribute to teachers leaving the profession, especially of those with attractive job prospects elsewhere.

Open Questions

There is little evidence on the effect of teachers' incentives at schools: will a reward structure that ties pay to teachers' performance remedy the problem? Is it feasible?

Three related questions:

- A. What are the principles that should guide the design of optimal teachers' incentives?
- B. Output versus effort?
- C. How much is performance or effort affected by creating incentives for current teachers?
- D. How much is performance or effort affected by changing the pool of teacher applicants?

Payment Based on Performance

Usually refers to some objective assessment of the teachers' effort• and/or success or on some measure of the performance of the students whom they teach

Advantages due to incentives and sorting/selection effects:

- It is fairer to reward teachers who perform well rather than paying all equally
- Performance-based pay motivates teachers and improves student performance
- A clearer connection between spending on schools and student performance builds public support
- Paying partly on the basis of performance may attract to teaching those who are best at improving performance

Limitations of Payment for Output

- Fair or accurate evaluation is difficult because performance cannot be determined objectively
- Co-operation among teachers is reduced; teachers are not motivated by financial reward
- Teaching becomes narrowly focused on the criteria being used; costs of implementation are too high
- Teachers might seek to game the rules: since they are subject to *multitasking*, they might concentrate on the easiest way to increase the regarded measure with little or no gains in the measure we care about
- Providing financial incentives may cause teachers to exert less effort because such schemes may demoralize them or harm their intrinsic motivation, particularly among teachers who as a group may have strong intrinsic motivation
- Some teachers, who previously exerted large effort, might decide to work less hard once they have reached their target expected income

Practical Concerns

- How to measure performance? Test Scores?
- Absolute or relative measures? Continuous or discrete?
- Should we care about distributional consideration?
- All versus some teachers, for example, subject-based **pay**
- Individual teachers' versus team contributions?
- Teams' incentives should be considered when production is a joint activity, is educational technology a team work?

Non-Experimental Evidence of Effect of PRP

- Only few countries introduced teachers' rewards that are related to reviewed performance and evidence of professional development
- Some countries (Denmark, England, Hungary, New Zealand, Norway, and Spain Sweden) provided an adjustment to the base salary for outstanding teaching performance
- Previous attempts at introducing performance-based reward programs have been poorly designed and implemented (Mohrman et al. 1996; Ramirez, 2001)
- Problems in developing fair and reliable indicator, and the training of evaluators to fairly apply these indicators have undermined attempts to implement programs (Storey, 2000)
- Poor goals clarity because of large number criteria, which restricts teachers' understanding of the program and makes implementation difficult (Richardson, 1999)

Non-Experimental Evidence of Effect of PRP

- Explanation of how, and on what criteria teachers are assessed may be difficult to articulate. When this occurs, it is almost impossible to give constructive feedback and maintain teacher support for the program (Chamberlin et al., 2002)
- Previous financial bonus has been comparatively small, which undermined the motivational value of the programs (Malen, 1999). When funding is limited, quotas are established, only a few teachers benefit and administrators find it difficult to explain why others missed out (Chamberlin et al., 2002).
- Other difficulties: opposition from teachers particularly related to concerns about unfair evaluation; and opposition from public school management systems to what are perceived as market based models.
- In the United States most attempts to implement merit pay for public school teachers have failed (Murnane, 1996).

Experimental Evidence of the Effect of PRP

- **Israel**, Two experiments, one based on team incentives (tournament among high schools in 1995-2001) and a second based on individual teachers' incentives (tournaments among math, English and Hebrew teachers in 2001). Lavy 2002 and 2004) show that:

- Individual incentives more effective than team incentives
- Incentives had large and significant effects on students performance
- No correlation b&e teachers' characteristics and performance/rewards
- More cost effective than other traditional interventions

- **England** , Performance-Related Pay, payment scheme based on pupil attainment. Atkinson et. al., (2005) show that:

- Individual incentives more effective than team incentives
- Scheme improved test scores and value added, on average by about half-a grade per pupil.
- Heterogeneity across subjects, maths teachers show no improvement-

Empirical Evidence of the Causal Effect of PRP

- **USA**, North Carolina's performance-compensation system, Clotfelter and Ladd (1996) and Johnson et. al., (1999) concluded that
 - bonuses had a positive effect as one element in a comprehensive accountability system, but were not viewed by teachers as major incentives.
- **Mexico**, since 1993, allowed teachers to gain permanent salary increase if they obtain scores above a cut-off in a national assessment of teachers. The awards are significant, 20% and more of the teachers' annual salary and they constitute a permanent increase. The assessment includes mostly measures of teachers' characteristics but it also includes an end of year students test scores as a criterion. McEwan and Santibañez (2005) conclude that the program led to
 - No gains in test scores of students

Empirical Evidence of the Causal Effect of PRP

- **Kenya**, randomly assigned 50 Kenyan primary schools to a treatment group eligible for monetary incentives (21-43% of monthly salary). The winning schools were determined by their test-score performance relative to other treated schools, and all teachers in such schools received awards. Glewwe, Ilias, and Kremer (2003) concluded that:

- The program produced test-score gains in treated schools, but they only lasted for the program's duration, and they appear to be the result of test preparation activities.

- **India**, In 60 informal one-teacher schools in rural India, randomly chosen, a financial incentive program was initiated to reduce absenteeism. A teacher's salary was a direct function of his attendance. Duflo and Hanna (2005) conclude that the program led to:

- An immediate decline in teacher absence by 25 percent
- Teachers were as likely to teach while in school-
- A a year later test scores in program schools were 0.17 standard-deviations higher than in the comparison schools and children were 40 percent more likely to be admitted into regular schools.

Other PRP Programs with no Evaluation

- Switzerland, Two cantons, Zurich (1999) and St. Gallen (2000) introduced links between teaching performance and pay:
 - Salary increments are provided over a period of years rather than applied on the basis of the assessment of a year's work
 - The assessment is undertaken by a team of representative of the school committee and it includes class observation, an interview with the teacher and the preparation by the teacher of a report describing his pedagogical approach.

No Evaluation of the effect of the program.

Other PRP Programs with no Evaluation

- **Sweden**, Introduced an individual teacher pay system in 1995 that includes elements of performance-based pay:

- Pay raises are linked to improved performance

- Principals can reward teacher if they work harder and take up more tasks than what is generally expected.

No Evaluation of the effect of the program

- **Finland**, The collective bargaining agreement for 2003/04 includes elements for assessing the standards of teacher's work:

- Provide the scope for municipalities to pay bonuses on the basis of individual professional proficiency and performance

- Introduce an individual teacher pay system in 1995 that includes -

Other PRP Programs with no Evaluation

- Chile has introduced in 1995 a system of merit awards to schools called the National System to Evaluate School Performance.
 - Schools are stratified so that competition is between relatively-comparable establishments
 - Separation of schools into homogenous groups makes the competition-more equitable
 - Ninety percent of the awards are paid directly to all the teachers as-bonuses. School directors may allocate the remaining 10 percent to outstanding teachers
 - Awards are based on an index composed of six factors: absolute levels-of student learning accounts for 37 percent of the index, changes in scores accounts for 28 percent, the remaining factors constituting 35 percent of the weight are also largely based on education outcomes.
- No evaluation of the effects of the program-

Evidence on Distortions Due to Accountability Measures

- Teacher cheating (e.g., Jacob & Levitt, 2003)
- Exclusion of low-achieving student from high-stake testing (e.g., Cullen & Reback, 2002; Figlio & Getzler, 2002)
- Test preparation and coaching (e.g., Glewwe, Ilias, & Kremer, 2003)
- Increasing students' caloric intake on the day of the test (Figlio & Winicki, 2005).

Conclusions

- Teachers may be expected to respond to incentives inherent in compensation structures
- General theories of compensation should apply to teaching
- Those theories suggest that output-based pay is best used when output is well defined and easily measured
- Input-based pay is best when jobs are inherently risky and when output is not easily observed
- The experimental evidence is encouraging in regard to whether teachers incentive are effective in schools but we are far from knowing what are the features of well-designed incentive system
- More experiments are needed
- Special attention to distortions created by incentives.