

Doing What Counts

Design Principles for a Study on Teacher Incentives

**Jason C. Snipes
Janet C. Quint
Shelley Rappaport
Lynne Steuerle Schofield**

**Prepared by MDRC
and the Laboratory for Student Success
at Temple University**



February 2006

This document is a product of MDRC and the Laboratory for Student Success (LSS) at Temple University Center for Research in Human Development and Education and has been supported, in part, by the Institute of Education Sciences of the U.S. Department of Education through a contract (ED-01-CO-0007) with LSS.

Inquiries about the work of LSS should be sent to LSS, Outreach and Dissemination Unit, Temple University, 1301 Cecil B. Moore Avenue, Philadelphia, PA 19122-6091, lss@temple.edu, or visit the website at <http://www.temple.edu/LSS>.

Dissemination of MDRC publications is supported by the following funders that help finance MDRC's public policy outreach and expanding efforts to communicate the results and implications of our work to policymakers, practitioners, and others: Alcoa Foundation, The Ambrose Monell Foundation, The Atlantic Philanthropies, Bristol-Myers Squibb Foundation, Open Society Institute, and The Starr Foundation. In addition, earnings from the MDRC Endowment help sustain our dissemination efforts. Contributors to the MDRC Endowment include Alcoa Foundation, The Ambrose Monell Foundation, Anheuser-Busch Foundation, Bristol-Myers Squibb Foundation, Charles Stewart Mott Foundation, Ford Foundation, The George Gund Foundation, The Grable Foundation, The Lizabeth and Frank Newman Charitable Foundation, The New York Times Company Foundation, Jan Nicholson, Paul H. O'Neill Charitable Foundation, John S. Reed, The Sandler Family Supporting Foundation, and The Stupski Family Fund, as well as other individual contributors.

Acknowledgments

This paper rests on the contributions of many people. First and foremost, we want to thank the district officials in the study sites, who helped us understand the details, successes, and challenges associated with the incentive plans in place in their district. While all of them gave generously of their time and insights, we want especially to cite the contributions of Brad Jupp at the Denver Public Schools and of Tomas Hanna, Sr., at the School District of Philadelphia.

Our fellow Doing What Counts team members were instrumental in designing the research, conducting interviews, and reviewing earlier drafts of the paper. At Temple University, Kent McGuire, Crystal Byndloss, Will Jordan, and Heidi Ramirez played critical roles in all these respects; so did Robert Ivry and Glee Holton at MDRC. Susan Wally also contributed to the research.

John Hutchins carefully edited the report, and Stephanie Cowell prepared it for publication.

Copyright © 2006 by MDRC and the Laboratory for Student Success, Temple University Center for Research in Human Development and Education. All rights reserved.

Contents

Introduction	1
Teacher Quality: A Literature Review	4
Literature on the Use of Incentives	12
Findings About Selected School Districts' Incentive Programs	17
Research Design for an Exploratory Study	26
Appendix: Potential Measures for Surveys of Teachers and Principals	44

Introduction

The most significant factor affecting students' learning in the classroom is the quality of their teachers.¹ Yet a central paradox of American education is that students from disadvantaged backgrounds, who arguably most need well-qualified and experienced teachers, are least likely to get such teachers. Research indicates both that the skills and backgrounds of teachers are among the most important determinants of students' academic success and that "teacher quality" is inequitably distributed across schools serving students from different racial and socioeconomic backgrounds. The Doing What Counts project is a collaborative effort of Temple University and MDRC² to design and execute rigorous research that will develop reliable evidence about the use of financial and non-financial incentives to recruit and retain high-quality teachers for underperforming schools — with the ultimate goal of improving student achievement.

For a variety of reasons, schools serving high proportions of disadvantaged and minority students present more challenging and less desirable contexts in which teachers must work. As Figure 1 shows, students often arrive at these schools behind their more advantaged peers in terms of academic achievement or even "school readiness," and they frequently face other problems that impede their academic progress. At the same time, these schools usually have fewer resources and confront a variety of other challenges. Teachers in these schools often feel that they must work harder than their counterparts in more advantaged schools to achieve the same (or even weaker) results. Subjected to increasing pressures for "accountability," teachers working in high-need schools frequently feel unsupported and unappreciated — "beaten up on" rather than thanked for taking on a difficult job.

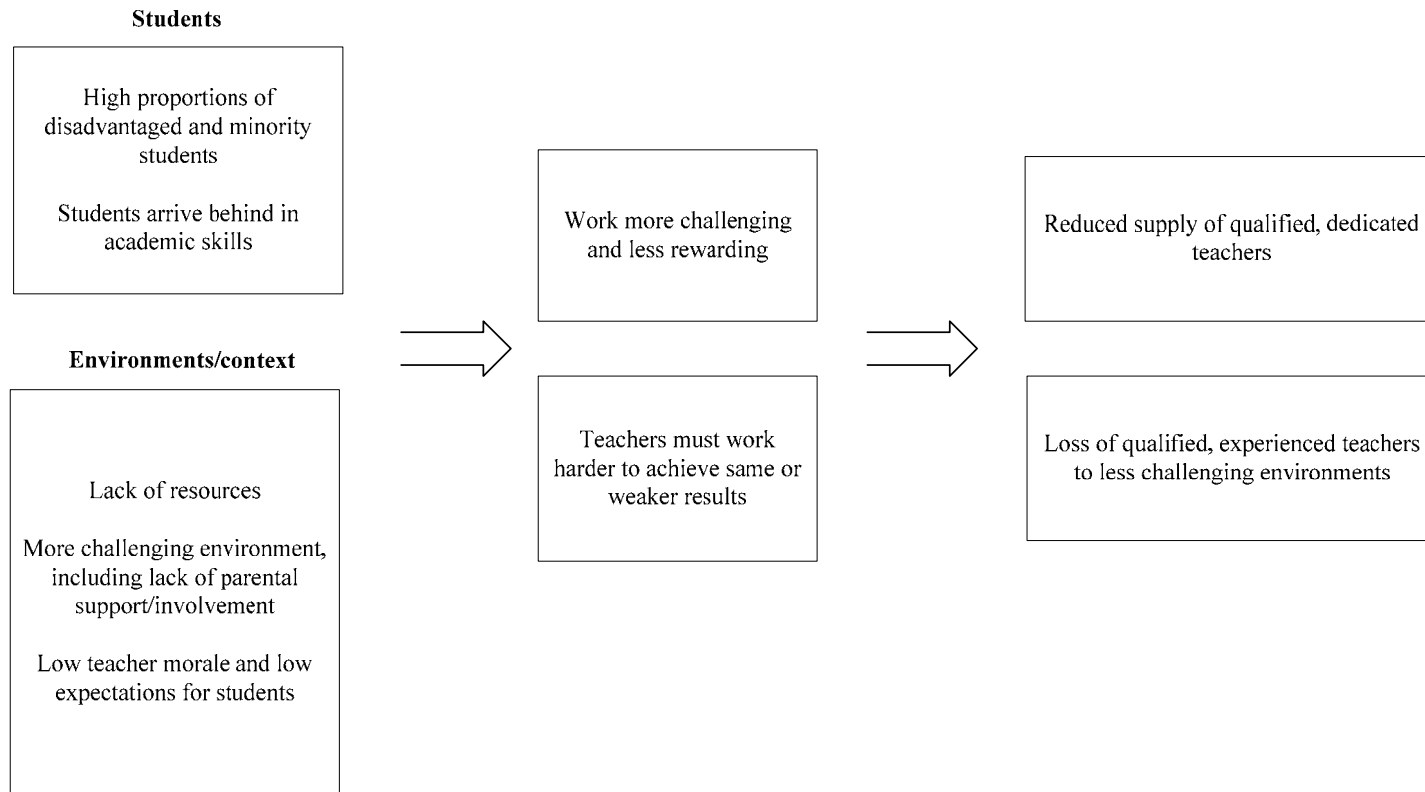
As a result, schools serving disadvantaged populations have more difficulty hiring and retaining highly qualified teachers. And as teachers in these schools gain experience, they frequently opt to move to less challenging contexts with more supportive and satisfying work environments — for the same or even more competitive salaries — leaving behind a teaching staff that is more junior and less skilled. Research has established that teachers in high-poverty schools with large proportions of nonwhite students typically are less credentialed and less experienced than their counterparts in more affluent schools, and turnover rates are typically high.

A number of urban school districts across the country have been investing in efforts to increase the supply of effective teachers in low-performing/high-need schools. Offering financial incentives and other inducements to teachers has been a key part of these efforts. The theory

¹National Commission on Teaching and America's Future. (1996). *What Matters Most: Teaching for America's Future*. New York: Columbia University.

²A representative from The Council of the Great City Schools is also providing support and input.

Figure 1
Challenges Facing Hard-to-Staff Schools



is that, by compensating teachers for some of the challenges they face in “high-need” environments, the incentives will result in an increase in the supply of teachers who are willing to work — and stay — in hard-to-staff schools.

There is reason to think that offering incentives to teachers may benefit students. Earlier MDRC work indicates that several large urban school districts that have succeeded in increasing student achievement have incorporated in their reform initiatives incentives for teachers with specific credentials, experience, and/or training to work in low-performing schools.³ However, there is no solid evidence about whether the incentive plans have actually increased the supply of high-quality teachers in these schools or increased student achievement.

The Doing What Counts project is designed to unfold in two stages. The first stage, which is described in this paper, will address a critical question:

- To what extent can well-designed incentive strategies change the supply of high-quality teachers in needy schools by changing recruitment and retention outcomes?

Positive results during this first stage would justify proceeding to a second stage, in which the MDRC-Temple team would examine a second question:

- To what extent do the changes in teacher supply brought about by incentive strategies affect the academic performance of students in these schools?

The primary purpose of this document is to set forth detailed design principles for the first stage of the study — an investigation of the effects of incentives on the recruitment and retention of high-quality teachers. First, however, the paper lays theoretical and research groundwork for the discussion of the study’s design. Section II presents a brief literature review about the characteristics that are (and are not) associated with teacher quality and describes how quality is measured. Section III describes how teachers are usually compensated in the United States and explains various kinds of incentives, financial and non-financial, that might be used to influence teacher behavior. Section IV recounts the results of a “reconnaissance mission” by MDRC-Temple team members to learn how teacher incentives have been used in seven urban school districts: Boston Public Schools, the Charlotte-Mecklenburg School District, Chattanooga Public Schools, Denver Public Schools, Los Angeles Unified School District, Miami-Dade County Public Schools, and the School District of Philadelphia. The team gathered information about these districts from the Internet, from published reports, and, most important, through telephone interviews with district officials. Because the Denver plan, known as Pro-

³Snipes, J., Doolittle, F., and Herlihy, C. (2002.) *Foundations for Success: Case Studies of How Urban School Systems Improve Student Achievement*. Washington, DC: Council of the Great City Schools.

Comp, is exceptionally well-developed, two team members made a site visit to Denver, where they were able to speak with a variety of individuals about the initiative.⁴

Finally, Section V, the core of the document, sets out the design principles for the first phase of the proposed research project. It addresses the outcomes of interest that should be examined and which method for measuring the effects of incentive strategies should be employed.

The discussion focuses primarily on financial incentives because this is where much policy interest has centered. A consistent theme of the document, however, is that non-financial incentives may be just as important as financial ones, if not more so. School districts have often put both kinds of incentives into place, so that an effort to evaluate district incentive strategies is likely to entail an examination of incentives that fall into both categories.

Teacher Quality: A Literature Review

This review briefly discusses three key aspects of the literature on teacher quality: the attributes that define high-quality teachers; the factors that determine teacher quality; and methods currently in use for identifying high-quality teachers. In general, the literature on these topics reveals considerable ambiguity and suggests a need for further investigation.

The Attributes of High-Quality Teachers

A number of groups have developed standards for defining high-quality teaching. In its book, *Testing Teacher Candidates: The Role of Licensure Tests in Improving Teacher Quality*,⁵ the National Research Council researchers compared the standards used by three prominent organizations: the National Board for Professional Teaching Standards (NBPTS), the Interstate New Teacher Assessment and Support Consortium (INTASC), and the National Council for Accreditation of Teacher Education (NCATE). All three organizations identified five attributes that define excellent teachers:

1. They are committed to their students and to students' learning.
2. They have deep subject matter knowledge.

⁴While in Denver, researchers met with the President of the Denver Public Schools Board of Education; the Executive Director of the Denver Classroom Teachers Association; the ProComp Transition Team; the Chief Information Officer, Chief Technology Officer, and Executive Director of Curriculum and Instruction for Denver Public Schools; the Senior Program Officer for the Rose Community Foundation; and a teacher and a principal.

⁵National Research Council Committee on Assessment and Teacher Quality. (2001). *Testing Teacher Candidates: The Role of Licensure Tests in Improving Teacher Quality*. Washington, DC: National Research Council.

3. They manage and monitor student learning.
4. They are reflective about their teaching.
5. They are members of a broader learning community.

Some studies have also attempted to identify special characteristics of effective teachers in urban settings. Sachs' review of the literature suggested that effective urban teachers should demonstrate five attributes: (1) sociocultural awareness; (2) contextual interpersonal skills; (3) self-understanding; (4) risk-taking; and (5) perceived efficacy.⁶ But when Sachs then surveyed a number of urban teachers, she found that these five attributes did *not* distinguish effective teacher from ineffective ones.

Martin Haberman, a professor emeritus at the University of Wisconsin–Milwaukee, developed the “Star Teacher” screener to help principals predict success among urban teachers. The screener asks teachers and potential teachers about their views and capacities in a number of areas: persistence, promoting learning, theory and practice, approach to “at-risk” students, professional and personal orientation to children, burnout, and fallibility.⁷ Evaluative studies of the screener have found that teachers hired using Haberman’s method perform at least as well as other teachers — *and* remain in the profession longer.⁸

In general, further research is needed to understand the specific attributes that make teachers successful in urban settings.

Determinants of Teacher Quality

A report issued by the National Council on Teacher Quality, *Increasing the Odds: How Good Policies Can Yield Better Teachers*, concluded that, among seven potential predictors of teacher quality, a teacher’s level of literacy and the selectivity of the college she or he attended are the best predictors of classroom effectiveness.⁹ (The other factors measured included possession of a master’s degree, years of experience, having taken education courses, holding traditional teaching certification, and subject area knowledge.) Similarly, studies by Ferguson, Ha-

⁶Sachs, S.K. (2004). Evaluation of teacher attributes as predictors of success in urban schools. *Journal of Teacher Education* 55(2): 177-187.

⁷Haberman, M. (2005). *National Center for Alternative Teacher Certification Information*, www.habermanfoundation.org. Retrieved October 21, 2005.

⁸Rothman, R. (2004). Landing the “highly qualified teacher”: How administrators can hire — and keep — the best. *Harvard Education Letter*. January/February 2004 issue.

⁹National Council on Teacher Quality. (2004). *Increasing the Odds: How Good Policies Can Yield Better Teachers*. New York: Author.

nushek, and Hedges and colleagues indicate that teachers' cognitive skills and their general level of academic achievement are strong predictors of their effectiveness with students.¹⁰

These findings suggest that it is important to recruit high academic achievers into the teaching profession, but they yield little information about what particular policies and practices prepare prospective teachers for the classroom. In fact, much remains to be learned about how to produce high-quality teachers. There is little agreement about what levels of content knowledge or pedagogical knowledge teacher preparation programs should prescribe, about the value of traditional vs. alternative teacher certification, and about how teaching experience affects the quality of instruction. Below is an account of the research on the value of subject area knowledge and pedagogical knowledge, the role of traditional certification vs. alternative certification, and the effect of teacher experience. Table 1 summarizes the findings.

Subject Area Knowledge and Pedagogical Knowledge

A 2003 literature review by the Education Commission of the States (ECS) concluded that there is “moderate support for the importance of subject matter knowledge ... and ... limited support that preparation in pedagogy can contribute significantly to effective teaching.”¹¹ While subject matter knowledge plays a key role with regard to teacher quality, the ECS review suggests that the law of diminishing returns may come into play with subject-specific classes — that is, teachers may reach a plateau, after which any additional classes they take no longer help them boost student achievement. Adding to this hypothesis is the lack of evidence that a master's degree adds much marginal value to a teacher's effectiveness. However, it should be noted that the ECS review did not examine what levels of subject-matter knowledge were needed to teach specific courses at particular grade levels. For instance, a third-grade teacher would obviously need to know different mathematics than a twelfth-grade calculus teacher.

The ECS study also suggests that, with respect to pedagogical coursework, classes in assessment and curriculum development and in classroom management may offer some marginal gain. However, the literature is unclear about the best way to obtain pedagogical skills — whether preparation should include more classes or longer clinical practice periods.

Traditional vs. Alternative Certification

There is a long-standing debate about the impact of type of certification — traditional vs. alternative — on teacher effectiveness. In large part, this stems from two philosophies about

¹⁰See, for instance, Hanushek E.A., and Rivkin, S.G. (2004). *How to Improve the Supply of High-Quality Teachers*. Brookings Paper on Education Policy. Washington, DC: Brookings Institution.

¹¹Education Commission of the States. (2003). *Eight Questions on Teacher Preparation: What Does the Research Say?* Denver, CO: Author.

Table 1
Summary of Research Literature on Six Attributes of Teacher Quality

Attribute	What We Know	What We Don't Know
Subject Area Knowledge	<ol style="list-style-type: none"> 1. Is important 2. May be a plateau effect after which additional content area knowledge is unimportant 	<ol style="list-style-type: none"> 1. Course effect 2. Grade-level effect
Pedagogical Knowledge	<ol style="list-style-type: none"> 1. Subject-specific courses are important 2. Assessment and instructional strategy classes are important 3. Classroom management classes are important 	<ol style="list-style-type: none"> 1. How best to impart knowledge 2. Types of clinical experiences that are best
Traditional Certification	<ol style="list-style-type: none"> 1. Many definitions 2. Gives teaching a sense of profession 	<ol style="list-style-type: none"> 1. What specific attributes of traditional certification produce the best results 2. Contexts in which traditional certification produces better results
Alternative Certification	<ol style="list-style-type: none"> 1. Even more definitions than traditional certification 2. Provides alternative pathways for teachers who cannot go through a traditional path 	<ol style="list-style-type: none"> 1. What specific attributes of alternative certification produce the best results 2. Contexts in which alternative certification produces better results
Teacher Experience	<ol style="list-style-type: none"> 1. Has an effect for the first 3-5 years 2. Currently is the most important factor in teacher pay 	<ol style="list-style-type: none"> 1. Effect of experience after first 3-5 years (results may be biased by percentage of "high-quality" teachers leaving sooner) 2. Best way to operationalize the true effect of teacher experience
Teacher Retention	<ol style="list-style-type: none"> 1. High rates of teacher attrition exist across the country, but especially in small private schools and large urban schools 2. Teachers cite leadership, safety, pay, and feeling unprepared as reasons they are leaving 	<ol style="list-style-type: none"> 1. How teacher attrition affects the structure and organization of schools 2. What needs to change in order for teachers to stay in teaching

how best to prepare teachers. Anne Turnbaugh Lockwood, the Issues Analysis Director at the American Association of School Administrators, has described the debate as being between “traditionalists” and “non-traditionalists”:

Traditionalists believe that high-quality teacher preparation is linked inextricably to completion of a program of teacher education that must be offered only by colleges and universities ... They maintain that increased programmatic rigor and heightened sets of standards for both programs and certification will result in a teacher workforce of higher quality ... Non-traditionalists argue for new paths into teaching to accommodate individuals with bachelor’s degrees but no coursework in education. Non-traditionalists also believe in “on the job” training provided by districts, with a strong mentorship component.¹²

Both sides can cite research examples to support their claims. However, when taken as a whole, the research literature does not offer much evidence that either path to certification generally produces better teachers.¹³ Also, few studies look at the specific *contexts* in which traditional or alternative certification might produce high-quality teachers. It is important to note that there are no clear definitions of either alternative or traditional certification. Since each state is responsible for determining its own standards for certifying teachers, it should not be surprising that there are many definitions of “traditional certification.” But alternative certification programs are even more varied across the states. These include such well-known programs as Teach for America, along with emergency certification.

Teaching Experience and Teacher Quality

Because teacher salaries are generally tied to years of teaching tenure, one might assume that experience has a direct effect on teaching quality. However, the research suggests that this may not be true. As Murnane and colleagues concluded:

The most commonly available characteristics, teacher education and experience, are important variables to consider, because they almost always enter into the determination of teacher pay. Thus it is plausible to think that they should be directly related to productivity...Yet they explain little of the variation in teacher effectiveness...¹⁴

¹²Lockwood, A.T. (2002). *Who Prepares Your Teachers? The Debate over Alternative Certification*. Arlington, VA: American Association of School Administrators.

¹³National Council on Teacher Quality (2004).

¹⁴Murnane, R., Singer, J., Willett, J., Kemple, J., and Olsen, R. (1991). *Who Will Teach? Policies That Matter*. Cambridge, MA: Harvard University Press.

The considerable research that has been conducted on teacher experience and retention suggests that a teacher's effectiveness may increase over her first three to five years on the job. After this initial period, however, there is little evidence that additional years of experience enhance teaching performance.¹⁵ But measuring the impact of experience is made difficult by the considerable amount of teacher turnover that currently exists. Urban school districts, in particular, show high rates of teacher attrition. For instance, Philadelphia reported losing approximately 23 percent of the teachers hired in 2002-2003 by the beginning of the 2003-2004 school year.¹⁶ Even more important, research has shown that teachers with higher scores on IQ tests and SATs are more likely to leave for other careers.¹⁷ Given that level of literacy and selectivity of college attended are the two best predictors of teacher quality, this finding suggests that, on average, more accomplished teachers are leaving the profession in greater numbers than their less accomplished counterparts.¹⁸

Richard Ingersoll has suggested another reason to identify and promote policies that curtail teacher attrition: While experience cannot be directly related to student achievement, teacher retention problems affect the ability of a school to function effectively. Ingersoll notes:

[H]igh rates of teacher turnover are of concern not only because they may be an indication of underlying problems in how well schools function, but also because they can be disruptive, in and of themselves, for the quality of school community and performance.¹⁹

Ingersoll's comment is a useful reminder about the importance of understanding how the structure and organization of schools can affect the quality of the teaching that takes place within them.

Identifying High-Quality Teachers

Since neither specific preparatory coursework nor type of certification is strongly predictive of high-quality teaching, other strategies for identifying effective teachers have been devised. Four such methods are described below. Two involve certification by independent professional boards, one calls for classroom observation, and one involves a statistical analysis of student performance.

¹⁵National Council on Teacher Quality (2004).

¹⁶Neild, R.C., Useem, E., and Farley, E. (2005). *The Quest for Quality: Recruiting and Retaining Teachers in Philadelphia*. Philadelphia: Research for Action.

¹⁷Hanushek and Rivkin (2004).

¹⁸This finding could help explain why, after five years, length of teaching tenure has little relationship to teacher quality.

¹⁹Ingersoll, R. (2001). *Teacher Turnover, Teacher Shortages, and the Organization of Schools*. Seattle, WA: Center for the Study of Teaching and Policy, University of Washington.

National Board for Professional Teaching Standards

The National Board for Professional Teaching Standards (NBPTS) has developed a method for identifying high-quality teachers that includes an extensive screening process. Candidates are required to submit portfolios of their assignments, student work samples, and videotapes of their teaching; these products receive thorough analysis. Teachers also complete a series of written exercises to demonstrate the depth of both their subject matter and pedagogical knowledge.²⁰ Many studies have shown that NBPTS certification has a high correlation with effective teaching, and many school districts provide incentives to teachers who go through the process and receive certification.

American Board for the Certification of Teacher Excellence

The American Board for the Certification of Teacher Excellence (ABCTE) was founded in 2001 to reduce barriers for people who want to enter the teaching profession. It awards two kinds of certification: Passport to Teaching and Master Teacher. The Passport to Teaching certification includes two exams, a Professional Teaching Knowledge Exam and a Subject Area Knowledge Exam. The Master Teacher certification also consists of two parts, a Subject Area Knowledge Exam and a demonstration of classroom effectiveness as determined by a longitudinal study of student academic achievement.²¹ Unfortunately, there has been little research on ABCTE certification and its link to student achievement over time, however such achievement is defined. As far as the MDRC-Temple team is aware, no districts currently offer incentives to teachers who choose to become Master Teachers through ABCTE.

Assessment of Teachers Using Rubrics of Instructional Quality

As an alternative to high-stakes testing, many groups — like the Center for Research on Evaluation, Standards, and Student Testing at UCLA and the Learning Research and Development Center at the University of Pittsburgh — are developing instruments to rate instructional quality, classroom practice, and teacher assignments. Through research grounded in cognitive psychology, these groups have attempted to determine what facets of teaching and classroom practice make up good instructional quality. They have then developed guidelines for scoring each of these aspects of practice. The motivation for developers of the scales is often the maxim

²⁰National Board for Professional Teaching Standards, <http://www.nbpts.org/about/index.cfm>. Retrieved April 19, 2005.

²¹American Board for the Certification of Teacher Excellence, <http://www.abcte.org/about.html>. Retrieved on April 19, 2005.

“what you test is what you get.”²² Although not necessarily a replacement for testing students, these education researchers believe that much can be learned by “testing” the teachers according to what they do on a daily basis inside the classroom.

Value-Added Models to Define Quality Teaching

When analyzed appropriately, standardized tests can be used to summarize the annual gains in student achievement, or the “value added” by each teacher. William Sanders of the University of Tennessee has developed the most recent version of value-added assessment, which employs an advanced form of analysis of variance (ANOVA).²³ Sanders argues that value-added models, which measure *changes* in achievement, can combat many of the problems associated with simply comparing the scores a particular teacher’s students achieve on standardized tests against a national average or standard. This decreases the problem of unfairly judging teachers in disadvantaged schools, whose students may exhibit considerable progress but still score significantly lower than the national average.²⁴ The results of value-added models have also been shown not to be affected by the confounding variables seen in other measures using standardized tests, such as race, socioeconomic status, and previous learning.²⁵

However, there are problems with value-added measures as well. The value-added models require annual testing of students in all grades with a reliable and valid achievement test.²⁶ Clearly, using tests with poor precision is problematic when trying to measure the added value of each teacher. Value-added models also require that the test items must be rewritten every year and connected to an underlying linear scale. Finally, value-added analysis cannot control for all of the confounding variables that could affect student scores, such as changes in income, changes in health, and changes in the environment.²⁷

* * *

²²Aschbacher, P.R. (1999). *Developing Indicators of Classroom Practice to Monitor and Support School Reform* (CSE Technical Report No. 513). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

²³Sanders, W.L., Saxton, A.M., and Horn, S.P. (1997). The Tennessee Value-Added Assessment System: A Quantitative, Outcomes-Based Approach to Educational Measurement. In Millman, J. (ed.), *Grading Teachers, Grading Schools: Is Student Achievement a Valid Evaluation Measure?* Thousand Oaks, CA: Corwin Press.

²⁴Stone, J.E. (1999). Value-Added Assessment: An Accountability Revolution. In Kanstoroomm, M., and Finn, C.E. (eds.), *Better Teachers, Better Schools*. Washington DC: Thomas B. Fordham Foundation.

²⁵University of Tennessee Value Added Research and Assessment Center. (1995). *Graphical Summary of Educational Findings from the Tennessee Value-Added Assessment System (TVAAS)*. Knoxville, TN: UT-VARC.

²⁶Stone (1999).

²⁷Stone (1999).

Clearly, there are many open questions associated with efforts both to produce and to identify high-quality teachers. The lack of consensus in the field suggests that the Doing What Counts study can make a large contribution in advancing the dialogue about what constitutes teacher quality, as well as in testing strategies to hire and retain high-quality teachers in high-need schools.

Literature on the Use of Incentives

The use of incentives to influence the work of employees is not a new concept in economics. The theory behind incentives is predicated on the basic assumption that workers will do more and better work if they believe they will receive appropriate rewards. In business, bonuses are often given to those individuals who contribute to increased profit margins or who expand the business into profitable areas. The basis for performance-based merit pay can be obvious: Did the business make more money? Are more consumers buying a product? Given the difficulty of measuring and defining teacher quality, however, the use of incentives in education is less straightforward. What measures should be used to reward teachers if a common definition of teacher quality is not agreed upon?

This section summarizes what is known about teacher compensation in general and about the use of two kinds of incentives — financial and non-financial — to reward desired teacher behaviors.

Teacher Compensation in the United States

Murnane and Cohen found that over 99 percent of teachers in the 1980s were paid using uniform schedules where salary increments were based on teaching experience and education (in the form of degrees awarded or credit hours amassed).²⁸ During that decade, calls for reform of the teacher compensation system began almost concurrently with calls for reform of the education system in general. The seminal 1983 report, *A Nation at Risk*, found that compensation systems generally rewarded teachers for staying in a particular school district for another year or for amassing more college credits. Teachers were not offered additional pay for quality teaching, for spending extra time in the classroom, or for improving student achievement. The report contained four recommendations about changes or improvements to the teacher compen-

²⁸Murnane R., and Cohen, D. (1985). *Merit Pay and the Evaluation Problem: Understanding Why Most Merit Plans Fail and a Few Survive*. Project Report No. 85-A14. Stanford, CA: Stanford Education Policy Institute, School of Education, Stanford University.

sation schedule, including one suggestion that teacher compensation be performance-based and another that incentives be used to recruit people into the profession.²⁹

Experience and credentials remain the basis for most compensation systems today. As one student of the subject concluded, “[T]eacher classroom performance or productivity is not considered part of salary evaluation with the fixed-step approach; hence there is no economic incentive for teachers to meet this important objective.”³⁰

Since 1983, both performance-based systems and recruiting incentives have been tried to varying degrees in some states. But, in general, these incentive systems have not been rigorously evaluated to determine their efficiency or effectiveness. Moreover, a survey by the National Association of State Boards of Education (NASBE) of state heads of teacher licensure across the country found that there was a disconnect between the teacher shortage areas identified by the states and the incentive programs offered by districts in these states.

Financial Incentives

The concept of performance-based pay or merit pay and incentives was not a new idea when it was suggested in *A Nation at Risk*. According to Troen and Boles, merit pay has been around since 1710, when it was used in England to reward teachers for increasing student scores on tests.³¹ Recent years have brought about a resurgence of the merit pay and incentive concepts. In almost all cases in the United States, merit pay has been used to supplement, not supplant, the current system. Monk and Jacobson designed a four-category typology of the performance criteria used to implement merit pay: (1) increased quantity of work; (2) increased level of efficiency; (3) increased teacher effectiveness or level of accomplishment; and (4) the importance of the accomplishment.³²

In reviewing the incentive systems operated by several school districts, MDRC-Temple team members found that most of the financial incentives fall under the following three categories:

- **Pay for Skills and Competencies.** Teachers are paid extra for having and/or developing a certain set of skills, with a variety of methods used to assess skills acquisition (e.g., achieving National Board for Professional Teaching

²⁹National Commission on Excellence in Education. (1983). *A Nation at Risk*. Washington, DC: U.S. Government Printing Office.

³⁰Bruno, J.E. (1986). Teacher compensation and incentive programs for large urban districts, *The Elementary School Journal*, 86(4): 424-447 (Special Issue: Policy Initiatives for Developing a Teaching Profession).

³¹Troen, V., and Boles, K.C. (2005, November 28). How ‘merit pay’ squelches teaching. *The Boston Globe*.

³²Jacobson, S.L. (1992). Performance related pay for teachers: The American experience. In Tomlinson, H., (ed.), *Performance Related Pay in Education* (pp. 34-54). London: Routledge.

Standards certification, completing a set of professional development courses).

- **Pay for Teaching in Difficult or Hard-to-Staff Positions.** Teachers are given a yearly bonus for teaching in certain schools — most often, those in urban or rural settings — where teacher attrition is a common problem. Teachers may also be paid extra for teaching in subject areas, like math or science, where a teacher shortage exists.
- **Pay for Producing Desired Results.** Teachers are paid extra when their students show academic improvement (typically measured using standardized tests). Teachers are also paid extra for satisfactory performance in the classroom (usually based on the results of formal evaluations conducted by principals).

These incentives are discussed in greater detail in the next section of this report.

A number of these incentive systems have been evaluated using correlational analysis, with inconclusive results. And vociferous debates about these systems have taken place. Some critics claim that financial incentives are ineffective. Some argue that merit is difficult to assess in teaching — and that, without a good measure of teacher performance, some good teachers will go unrewarded while other teachers will receive undeserving rewards.³³ Others claim that merit-pay incentives work on the basis of competition, while teaching is a profession based on cooperation.³⁴ Still others argue that teachers' unions do not support merit-pay incentives and that, without union support, no new teacher compensation system will be well-implemented. Finally, some contend that people who enter the teaching profession are driven less by monetary incentives than by altruistic motives.³⁵

There are also studies that suggest that incentives work to achieve certain specific objectives. Johnson indicates that “[while] financial incentives can promote specific behaviors (such as taking on difficult teaching assignments) and can direct teachers’ efforts to measurable goals (such as achieving higher test scores), they are less promising as tools to improve general teaching performance.”³⁶ However, it is clear that there is much work left to be done to describe and understand how financial incentives can be used to recruit and retain high-quality teachers, as well as to improve teaching quality more broadly.

³³Bruno (1986).

³⁴Jacobson, S.L. (1995). Monetary incentives and the reform of teacher compensation: A persistent organizational dilemma. *International Journal of Educational Reform*, 4(1): 29-35.

³⁵See Stern, D. (1986). Compensation for teachers. *Review of Research in Education*, 13: 285-316; Jacobson (1995).

³⁶Johnson, S.M. (1986). Incentives for teachers: What motivates, what matters. *Educational Administration Quarterly* 22: 54-79.

Non-Financial Incentives

Financial incentives may not be enough to retain and attract teachers to low-performing schools. Ozcan concluded that teachers tend to downplay the importance of economic rewards and to overemphasize the intrinsic awards.³⁷ A study using National Center for Education Statistics data used regression analysis to compare the relative importance of various factors in explaining teacher satisfaction. The study found that administrative support and leadership, student behavior and school atmosphere, and teacher control over the working environment were all more closely associated with teacher satisfaction than were variables related to compensation.³⁸

Poor working conditions — among them, lack of basic materials and time, isolation, assignment to non-essential duties, little input into the design and organization of schools, and limited opportunities for career advancement and professional growth and support — are closely related to difficulty in recruiting and retaining teachers.³⁹ Satisfaction among teachers increases when they receive support from administrators, cooperation from their colleagues, and the resources they need, and when they are not burdened with non-teaching duties. Teacher satisfaction is also higher in schools where student apathy, misbehavior, and violence do not present problems.⁴⁰ Teachers report moving to schools with better student discipline and smaller class sizes.⁴¹

Two kinds of non-financial incentives — opportunities for professional advancement and recognition and opportunities for professional development and support — offer promise for retaining high-quality teachers.

Opportunities for Professional Advancement and Recognition

Richard Ingersoll found that it takes not only higher wages but also opportunities for professional advancement to retain teachers.⁴² Similarly, Dorman and Fulford's report, *Teacher Incentives*

³⁷Ozcan, M. (1996, April 8-12). *Improving Teacher Performance: Toward a Theory of Teacher Motivation*. Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY.

³⁸Each of the sets of variables related to working conditions explained between 0.14 and 0.17 of the total variance in teacher satisfaction, whereas variables related to compensation explained only 0.07 of that variation. See Perie, M., and Baker, D.P. (1997). *Job Satisfaction Among America's Teachers: Effects of Workplace Conditions, Background Characteristics, and Teacher Compensation*. Report for the U.S. Department of Education. Washington, DC: American Institutes for Research.

³⁹Hirsch, E. (2004). *Listening to the Experts: A Report on the 2004 South Carolina Teacher Working Conditions Survey*. Chapel Hill, NC: Southeast Center for Teaching Quality.

⁴⁰Perie and Baker (1997).

⁴¹Ingersoll, R. (2003). *Is There Really a Teacher Shortage?* Seattle, WA: Center for the Study of Teaching and Policy.

⁴²Ingersoll (2003).

from the Inside: Five Studies by Teacher Researchers, concluded that increased professional opportunities lead to increased teacher motivation and satisfaction.⁴³ For many teachers, the only path to advancement is to move out of the classroom into administrative positions. The Teaching Commission recommends establishing career-advancement paths that allow accomplished teachers to stay in the classroom while serving as mentors to their less-experienced peers.⁴⁴

Opportunities for Professional Development and Support

The new teacher attrition rate is estimated to be between 12 to 20 percent in the first year of teaching⁴⁵ and nearly 40 percent after four years.⁴⁶ When new teachers leave the system, they take with them the knowledge and experience they have acquired at great financial cost to school systems.⁴⁷ Comprehensive induction programs — which include release time, full-time mentors and coaches, and opportunities for networking — can help give new teachers the support they need as they enter the profession.⁴⁸ Low-income, high-minority, and linguistically diverse districts have been found to be the most likely to benefit from these programs.⁴⁹

Professional development may have a positive effect on the retention of more seasoned teachers as well. A report from the Southeast Center for Teaching Quality found a significant connection between retention and professional development. At the elementary level, professional development was found to exert a significant positive effect on teacher retention — but, at the high school level, the connection was negative. The report's authors speculate that this negative finding was due to the unique nature and small sample (96) of high schools.⁵⁰

The recognition that teachers respond to non-financial incentives as well as financial ones has driven the creation of organizations such as Resources for Indispensable Schools and Educators (RISE) and the Milken Foundation's Teacher Advancement Program (TAP), both

⁴³Dorman, A., and Fulford, N. (1990). *Teacher Incentives from the Inside: Five Studies by Teacher Researchers*. Elmhurst, IL: North Central Regional Educational Laboratory.

⁴⁴The Teaching Commission. (2004). *Teaching at Risk: A Call to Action*. New York: The Teaching Commission, CUNY Graduate Center.

⁴⁵Cornett, L. (2002). *Quality Teaching: Can Incentive Policies Make a Difference?* Atlanta, GA: Southern Regional Education Board.

⁴⁶Moir, E. (2003, July 12-14). *Launching the Next Generation of Teachers Through Quality Induction*. Paper presented at the State Partners Symposium of the National Commission on Teaching & America's Future, Denver, CO.

⁴⁷Moir (2003).

⁴⁸Hirsch, E. (2004). *Listening to the Experts: A Report on the 2004 South Carolina Teacher Working Conditions Survey*. Chapel Hill, NC: Southeast Center for Teaching Quality.

⁴⁹Moir (2003).

⁵⁰Hirsch, E. (2004). *Teacher Working Conditions are Student Learning Conditions: A Report to Governor Mike Easley on the 2004 North Carolina Teacher Working Conditions Survey*. Chapel Hill, NC: Southeast Center for Teaching Quality.

based in California. RISE aims to stop the attrition of effective teachers by putting those committed to working in low-income schools into better teaching environments. Among the benefits that RISE offers are a support group for teachers, career opportunities, a community of like-minded educators, and financial rewards to defray the cost of materials and supplies.⁵¹ TAP helps ensure that highly skilled and strongly motivated people enter the teaching profession by offering them sustained opportunities for career advancement, ongoing school-based professional development, instructionally focused accountability, and performance pay.⁵²

As the next section indicates, financial and non-financial incentives are being tried in a number of places across the country. Research suggests that some combination of the two will be necessary to attract high-quality teachers to low-performing schools.

Findings About Selected School Districts' Incentive Programs

To better understand the use of incentives and to inform the project's research design, the MDRC-Temple team examined the incentive plans that six large urban school districts have undertaken. These districts were ones that had been identified as operating especially interesting and/or innovative incentive programs: the Charlotte-Mecklenburg School District, Chattanooga Public Schools, Denver Public Schools, Los Angeles Unified School District, Miami-Dade County Public Schools, and the School District of Philadelphia.⁵³ In addition to gathering information from the Internet and published reports, the team conducted telephone interviews with officials in the human resources department in each district. Team members also made a site visit to Denver and met with a variety of individuals associated with the design and implementation of that district's incentive program, known as ProComp.

This section reports the findings of this reconnaissance effort. It begins with a general description of the incentive programs that were studied, considers the financial and non-financial incentives offered by the districts and the extent of variation within each kind of incentive, and concludes with a discussion of overarching lessons.

General Findings

The financial incentive programs examined fit into three broad categories:

- Pay for teaching in difficult or hard-to-staff positions

⁵¹RISE website. Retrieved from <http://www.risenetwork.org> on October 24, 2005.

⁵²TAP website. Retrieved from <http://www.tapschools.org> on October 25, 2005.

⁵³In Miami-Dade, the team only looked at the incentive program offered in schools in the School Improvement Zone, known as the "Zone," which includes the 39 lowest-performing schools in the district.

- Pay for skills and competencies
- Pay for producing desired results

In addition, various non-financial incentives are used to attract teachers in the districts.

Table 2, which summarizes the findings, indicates that all six districts have incentives in place or planned for the near future for teaching in difficult or hard-to-staff positions. Five districts include incentives for having or acquiring skills and competencies, and four of the five have specific incentives for teachers who achieve certification by the National Board for Professional Teaching Standards. Three of the districts pay schools, teachers, or both for producing desired results. There is considerable variation among the districts in how they implement each category of financial incentive — for instance, in how they define who should be entitled to receive an incentive and how large the incentive should be.

In addition, districts differ in the variety of financial incentives they offer. For instance, while Denver and Charlotte-Mecklenburg offer incentives across all the categories, Miami-Dade has pooled all its resources into one incentive program. They also vary in how the incentives operate. In Denver, most of the individual incentives are quite small, but teachers can earn as much as \$500,000 in additional compensation over the course of their careers through what was described by one interviewee as a “Lego-like” approach to adding to their salary and qualifying for bonuses. For example, a middle school math teacher who succeeds in increasing the scores of English language learners in a hard-to-staff school would be eligible for one salary increase and two bonuses. In contrast, in Miami-Dade, the incentive immediately increases teachers’ salaries by 20 percent.

Finally, although Table 2 organizes all the specific incentives as fitting neatly in one category or another, many of the district programs reach across these categories. This is the case, for example, in districts that offer incentives to teachers for developing their skills only if they teach in hard-to-staff schools.

Pay for Teaching in Difficult or Hard-to-Staff Schools and Positions

Four of the districts studied had incentives to attract higher-quality teachers to certain schools. However, the criteria used to identify which schools are tied to more money for teachers differ. For instance, in Denver, the criteria used to identify a school as hard-to-serve includes the proportion of students receiving free and reduced priced lunch (FRPL), the percentage of students eligible for Medicaid, the percentage of English language learners, the percentage identified as special education, and the crime level of the neighborhoods in which students live. (The last criterion is included as a proxy for undocumented students that might not receive FRPL or

Table 2

Type	Sites	What Incentives Look Like
Pay for Teaching in Difficult or Hard-to-Staff Positions		
Pay for Teaching in Hard-to-Staff Schools	<p>Charlotte-Mecklenburg, NC</p> <p>Denver, CO</p> <p>Los Angeles, CA</p> <p>Miami-Dade, FL</p>	<ul style="list-style-type: none"> • A \$3,000 differential signing bonus for teaching in a high-need school (instead of the usual \$1,000 for a non-high need school) • Annual incentive for master teachers (\$2,500 for teachers with master’s degrees [MAs] and four years teaching, proficient student achievement, or other criteria like National Board for Professional Teaching Standards [NBPTS] and \$1,500 for teachers pursuing MA and eight years teaching) • 3% of Index Incentive Pay • Urban Classroom Teacher Program (UCTP) — \$1,000 in certain low-performing schools • 20% more pay for 20% more work in School Improvement Zone
Pay for Teaching in Shortage Areas	<p>Charlotte-Mecklenburg, NC</p> <p>Denver, CO</p> <p>Philadelphia, PA</p> <p>Los Angeles</p>	<ul style="list-style-type: none"> • One-time bonus of \$500 • 3% of Index Incentive Pay • \$1500/year • Up to \$3,000 for new bilingual teachers depending on assignment
Incentives for Developing Skills & Competencies		
Knowledge & Skill-Based Pay (not including pay for National Board certification)	<p>Charlotte-Mecklenburg, NC</p> <p>Chattanooga, TN</p> <p>Denver, CO</p>	<ul style="list-style-type: none"> • Free or reduced tuition for teachers in high-need schools pursuing MA or pedagogy courses towards regular certification. • Free MA in high-need schools • 9% of salary index for graduate degree • Tuition reimbursement up to \$1,000 • 2% of salary index per unit for participation in an ongoing course of professional development

(continued)

Table 2 (continued)

Type	Sites	What Incentives Look Like
Pay for National Board Certification	<p>Philadelphia, PA</p> <p>Charlotte-Mecklenburg, NC</p> <p>Denver, CO</p> <p>Philadelphia, PA</p> <p>Los Angeles, CA</p>	<ul style="list-style-type: none"> • Tuition reimbursement for a master's degree up to six credits (\$2,400) if in Incentive School • 12% higher salary schedule • 9% of salary index once certified • Reimbursed for fee and expenses up to \$2,500 once passed plus one day off during candidacy • \$3,500 for length of certificate in return for mentoring NBPTS candidates • 15% salary increase (plus CA offers a \$20,000 bonus over four years if NBCT works in low-performing school)
Pay for Producing Desired Results	Charlotte-Mecklenburg, NC	<ul style="list-style-type: none"> • Local Accountability Bonus: \$500-\$850 for all staff when students meet performance growth goals for AYP and/or NC ABCs (see below) • Deferred Supplemental Local Accountability Bonus: for all staff members in high-need schools that meet the annual student performance criteria. The bonus payment was held in escrow until the employee completed two additional years after the year the bonus was earned at the same school. Note: This deferred supplemental bonus plan is being discontinued because it has been seen as ineffective. • State ABCs Award: Pay for all certified staff and teaching assistants (\$750 certified, \$375 non-certified) for meeting ABC goals
Teacher Level	<p>Denver, CO</p> <p>Charlotte-Mecklenburg, NC</p> <p>Chattanooga, TN</p> <p>Denver, CO</p>	<ul style="list-style-type: none"> • 2% index for serving in a Distinguished School (schools identified based on student growth and other measures) • Bonus for full-time licensed or unlicensed staff for student achievement gains on standardized tests (\$2,000 elementary, \$5,000 high school) • Annual \$5,000 bonus for teachers if students show at least 1.25% annual growth in literacy skills on the Terra Nova • 1% of Index Salary when probationary teachers rated satisfactory on annual teacher evaluation • 3% of Index Salary when non-probationary teachers rated satisfactory on teacher

Table 2 (continued)

Type	Sites	What Incentives Look Like
		<ul style="list-style-type: none"> • evaluation every three years • 1% of Index Salary Pay (added to permanent salary) for meeting both student growth objectives • 1% of Index Incentive Pay (one-time bonus of \$333) for meeting one student growth objective • 3% of Index sustainable increase for significantly exceeding expectations. • Sustainable decrease of 3% of index for falling significantly below expectations only if previous increase earned (based on Colorado Student Assessment Program expectations).
Non-Financial Incentives Professional Development and Support	<p>Charlotte-Mecklenburg, NC</p> <p>Chattanooga, TN</p> <p>Philadelphia, PA</p>	<ul style="list-style-type: none"> • NBPTS support system for those seeking certification • Extensive professional development and leadership training in Benwood Schools (not clear if this is for teachers or principals or both) • Teachers at targeted incentive schools provided with “targeted professional development” dealing with managing disruptive pupil behavior • Coaches for every first-year teacher (coaches are discipline-based, not school-based) • Common, district-wide curriculum with a recommended pacing schedule and PD aligned with it
School Environment	<p>Charlotte-Mecklenburg, NC</p> <p>Chattanooga, TN</p> <p>Miami-Dade, FL</p> <p>Los Angeles, CA</p>	<ul style="list-style-type: none"> • Reduced class size in early grades in high-need schools • NBCTs being deployed in clusters to inspire professional learning community • Six out of nine principals replaced, and the ones in schools now get leadership training • National Board learning communities • Plan in the works to ensure every school has an effective principal

Medicaid but are income-eligible to do so.) Although there is about a 75 percent overlap between those schools that are low-performing and those that are identified as hard-to-serve, at least two low-performing schools did not qualify as hard-to-serve. In contrast, incentive programs in Miami-Dade, Chattanooga, and Los Angeles are targeted at the lowest-performing schools. In Charlotte-Mecklenburg, incentives are targeted at schools that are high-poverty as well as low-performance, while, in Philadelphia, the new criteria for a school to receive an incentive have not yet been set.

While there is a great deal of overlap between schools that have disadvantaged populations and those that are low performers, Denver's ProComp plan deliberately distinguishes between the two. The rationale for this distinction is to avoid creating an incentive for a school to remain on the low-performing list — but, rather, to create an incentive for a school to reach a high level of performance and stay there (even while it remains disadvantaged). Similarly, teachers in Miami-Dade's "Zone" need not worry about losing their incentive if schools improve. Zone schools remain in the Zone for the duration of the three-year project.

The amount of the incentives given to teachers for working in particular schools also varies across districts. On one end of the pay spectrum, teachers who participate in the Urban Classroom Teacher Program in Los Angeles get an extra \$1,000 for working in low-performing schools. According to one interview respondent, this program is seen as ineffective because the amount of money is not enough to attract teachers to incentive schools. At the other end of the continuum, teachers in Miami-Dade who opt to work in the Zone get 20 percent higher pay than teachers in schools outside of the Zone. Following implementation of the incentive program, the district ended up with a younger cadre of teachers on average who also have more experience, more teaching credentials (including National Board Certification), and more content-specific knowledge. In addition, by the time the project was six months old, many of the Zone's goals had already been accomplished. The leadership team set out to improve the rating of ten schools and it was able to do so in 15. Also, an unprecedented number of Miami-Dade County Public Schools children in the Zone were able to move up in reading level.

Pay for Skills and Competencies

Another strategy for ensuring that schools get the teachers they need is by either rewarding teachers who already possess desired skills and competencies or by home-growing those teachers. Four of the districts examined seem to have had success offering teachers financial incentives for having National Board Certification. In Denver, Charlotte-Mecklenburg, and Los Angeles, the incentives are equal to 9 percent, 12 percent, and 15 percent of teachers' salary, respectively. In Philadelphia, teachers are reimbursed for the \$2,300 National Board fee if they pass, and they are entitled to a day off during candidacy. Once they receive certification, they also get \$3,500 annually for the length of the certificate (10 years) in return for mentoring

new candidates for certification. In none of these cases is a teacher required to work in a high-need school to receive the incentive. However, in California, National Board Certified teachers are paid \$20,000 over the course of four years to work in low-performing schools. Compared to other cities with large numbers of National Board Certified teachers, Los Angeles does the best job of distributing them across low-performing schools, and California is the only state that has an equitable distribution of National Board Certified Teachers.⁵⁴

Districts can also seek to promote quality by offering teachers opportunities to develop their skills through degree and certification programs and professional development. In Chattanooga, teachers were asked to identify what it would take for all third-graders in the lowest-performing schools to attain at least grade-level reading skills and to get good teachers to come and stay in these schools. Their answer: more and better learning opportunities for teachers. In response, teachers in the lowest-performing schools are offered a free master's degree focusing on literacy instruction and urban teaching in exchange for staying in the same school four years after completing the degree and committing to serve as teacher-leaders in their school's instructional improvement efforts. This initiative is largely seen as a successful one, and it has now been extended to the five middle schools that feed into the lowest performing elementary schools. Of the four districts that offer a free or reduced-price degree or tuition towards university courses, three target the incentive to teachers in high-need schools.

Pay for Producing Desired Results

Schools are under pressure to increase their math and reading scores. It therefore makes sense to reward teachers who directly help increase those scores, but the issue of who should be paid an incentive when students show academic growth is a difficult one. The argument is made that all teachers in a school work together to teach each student and, therefore, all teachers who work with students who show improvement should be rewarded. Although students are not typically given standardized tests in kindergarten and first grade, the amount of work that these early childhood teachers do to prepare their students for when they do have to be tested affects their future achievement. In the same vein, physical education teachers help students remain healthy and build self-esteem, which also can impact how well students do academically.

Charlotte-Mecklenburg addresses this issue with its Local Accountability Bonus program, in which all staff members at all schools that meet annual student performance criteria receive incentive pay. This award has been popular with school staff and the general public. In

⁵⁴Wayne, A., Chang-Ross, C., Daniels, M., Knowles, K., Mitchell, K., and Price, T. (2004). *Exploring Differences in Minority and Majority Teachers' Decisions about and Preparation for NBPTS Certification*. Arlington, VA: SRI International; Humphrey, D.C., Koppich, J., and Hough, H.J. (2005, March 3). Sharing the wealth: National Board certified teachers and the students who need them most. *Education Policy Analysis Archives*, 13(18).

addition, the teachers interviewed by the MDRC-Temple team generally believe that this bonus is the most effective of their various incentive programs. In the first few years after these bonuses were introduced, there was a dramatic increase in student achievement; the increases have not been so impressive since. Nonetheless, last year the state legislature had to scramble to come up with funding for the state awards because more schools were eligible than had been anticipated.

Denver has dealt with this same issue by creating opportunities for teachers in all subject areas to write student growth objectives, which they then have to meet in order to receive an award. The pilot phase of Denver's incentive plan (the Pay-for-Performance pilot) was evaluated by the Community Training and Assistance Center (CTAC). Results showed that students whose teachers developed the highest quality objectives, based on a rubric developed by CTAC, averaged greater gains in achievement on a standardized test (ITBS), whether the teachers' objectives were met or not, than students whose teacher objectives were scored lower on the rubric.⁵⁵

Of course, rewarding every person who works in a school is quite costly. Perhaps this is why individual teachers in Chattanooga can earn \$5,000 if their students show at least 1.25 percent annual growth in literacy skills on the Terra Nova, while the Local Accountability Bonus in Charlotte-Mecklenburg only pays between \$500-\$850 to each staff member in schools meeting performance growth goals.

Non-Financial Considerations

Financial incentives may not be enough to retain and attract teachers to low-performing schools, according to some district personnel interviewed for this project. The School District of Philadelphia has been battling to recruit and retain certified teachers by advertising the financial rewards of working in the district and trying to dispel rumors that Philadelphia is a dangerous and unappealing place to work. The district also implemented a common, district-wide curriculum with a recommended pacing schedule, curriculum frameworks, and standardized professional development, as well as six-week benchmark tests to help keep teachers and students on course. A rationale for implementing this curriculum was that it would give new teachers a road map to follow that would reduce their anxiety and allow them to better focus on other issues in teaching and learning, while letting them feel a part of the larger teacher community. In addition, new teachers are given coaches. Something is working in Philadelphia — last year, 93 percent of new teachers remained with the district, in comparison with a retention rate of only 73 percent for new teachers hired during the 1999-2000 school year — although it is not clear what has caused this change.

⁵⁵Denver Classroom Teachers Association. (2005). *Straight Talk about ProComp: The Professional Compensation System for Teachers*. Denver: Author. Available at www.denverprocomp.org.

Having supportive principals is also important for recruiting and retaining teachers. In Los Angeles, the Human Resources Department is working on a plan to ensure that every school has an effective principal. In Chattanooga, part of the effort to improve achievement included replacing six out of nine principals in low-performing schools and providing leadership training.

Other Lessons Learned

Although these incentive programs have not been formally evaluated, the district staff members shared their beliefs about the effects that the programs have had. In general, incentives considered successful by interviewees were those that appeared accessible to a majority of people and those that paid a sufficient amount of money.

Accessibility

Denver respondents maintained that for an incentive to be effective, it has to gain widespread support. Most teachers fit into the average/adequate quality range, so an incentive that is going to enjoy significant buy-in from teachers has to appeal to teachers in this category. Denver's incentive plan was designed to reward average teachers who show continuous improvement as well as those that are exceptional, have sought-after skills, and/or choose to teach in hard-to-staff positions or schools. Since most teachers can potentially benefit from Denver's ProComp, it has widespread appeal — 59 percent of the members of the Denver Classroom Teachers Association voted to approve the compensation plan.⁵⁶

In Charlotte-Mecklenburg, a popular incentive program believed to be the most effective is the one that rewards the whole staff (including janitors) when students show student growth.

Adequacy

When it comes to incentives, district officials who were interviewed in Charlotte-Mecklenburg reported that anything “under \$5,000 doesn't have much impact.” The district had an incentive program called the Deferred Supplemental Local Accountability Bonus, which paid less than \$1,000 to each staff member in high-need schools that met the annual student performance criteria. The bonus payment was held in escrow until the employee completed two additional years at the same school. The reason this deferred supplemental bonus plan is being discontinued is because it has been seen as ineffective. According to interviewees, the amount was too small to affect retention. Another small incentive program is Los Angeles' Urban Classroom Teacher Program, which pays \$1,000 to each teacher in particular low-performing

⁵⁶Denver Public Schools and Denver Classroom Teachers Association. (2004). *ProComp: Denver Public Schools Professional Compensation System for Teachers*. Denver: Authors. Available at: <http://www.dpsk12.org>.

schools. An official in the district's Human Resource Department asserted that \$1,000 is not enough to make a difference.

The districts studied do have some incentives that are large enough to have the potential to make a difference. Teachers in Miami-Dade's School Improvement Zone get 20 percent more money than teachers in schools outside of the Zone. Teachers in Charlotte-Mecklenburg's and Chattanooga's high-need schools can get free or radically reduced tuition while pursuing their master's degrees. Once they receive their degrees, teachers in high-need schools are paid an additional \$2,500 per year. (However, some principals in Chattanooga have said that these incentives do not necessarily reward the best performers and that the criteria do not ensure that teachers will do well in a high-need school.)

In Chattanooga, too, the city's mayor and business leaders launched a differentiated pay program that provides \$5,000 annually to teachers whose students (tested on the Terra Nova) show at least 1.25 percent annual growth in literacy skills. In Chattanooga, the results have been dramatic in the lowest-performing schools: in 2002, the district had to fill 106 teaching vacancies in these 9 urban schools; in 2004, they had to fill only 28. As teaching quality stabilized in these schools, the achievement of students made very strong gains. While all Hamilton County schools made strides forward in student achievement, the gains made by Benwood schools (the lowest-performing schools), as measured by third-graders' reading results, surpassed the district's average gain among all third-graders, including those from better-resourced families and schools.

* * *

The results of the team's reconnaissance efforts indicate that while different kinds of financial and non-financial incentives are conceptually distinct, they are often linked in districts' compensation plans. This finding suggests that an evaluation aimed at assessing the relationship between incentives and teacher recruitment and retention outcomes is likely to test a *package* of incentives of various types. A careful study of this relationship will involve both qualitative and quantitative research aimed at understanding how these incentives interact to affect teachers' decisions about where and whom they will teach.

Research Design for an Exploratory Study

The ultimate goal of the line of research discussed in this paper is to ascertain the effects of teacher incentives on student achievement. But before undertaking a large-scale demonstration to test these effects, it is important to gather additional empirical evidence on the relationship between incentives designed to recruit teachers into typically hard-to-staff schools and actual changes in recruitment and retention outcomes. This information would help researchers assess the likelihood that an intervention involving incentives would increase the supply of

high-quality teachers in these schools, and that a large-scale random assignment evaluation of the impact of these teachers on student achievement would therefore be a worthwhile investment. An exploratory study would also help to develop and refine the intervention that would be studied. For example, some of the literature discussed earlier suggests that, rather than financial rewards, the non-pecuniary aspects of their jobs are the most important aspects of teachers' motivation. To the extent this is true, an exploratory study might find little association between financial incentives and changes in the recruitment and retention of qualified teachers in high-need schools. Such a finding would call into question the desirability of an expensive large-scale demonstration. It might also suggest that financial incentives were only associated with changes in teacher supply or average teacher characteristics when combined with other efforts to improve conditions or support teaching and learning in the targeted schools.

This section describes the design for this exploratory study, which is organized around three broad tasks:

- Understanding the local contexts in which incentive programs are implemented;
- Analyzing the implementation of incentive initiatives and how the initiatives interact with the local context; and
- Measuring the empirical connections between implementation of the incentive programs and changes in teacher recruitment and retention outcomes, including “teacher quality” in hard-to-staff schools.

It is important to emphasize that this exploratory study will yield suggestive rather than definitive findings. The design focuses on districts that have already implemented incentive systems for recruiting teachers to districts' “high-need” or “hard-to-staff” schools.⁵⁷ Because these systems are already in place, changing the way that incentives are allocated in order to randomly assign schools to different incentive regimes is not likely to be a realistic option. For this reason, a random assignment evaluation design — considered the “gold standard” of evaluation research — is not feasible.

Instead, the design employs an interrupted time-series design, along with comparison schools, to estimate the effects of the incentive programs on teacher recruitment and retention outcomes. In this design, changes in the size and qualifications of the teaching staff in high-need

⁵⁷An alternative would be to design an optimal incentive plan and then recruit districts that do not currently offer incentives to put this plan into effect. This would be a long, challenging, and expensive process, however, and it is unclear at this point how the incentives would be financed. A better course is to test the effectiveness of existing incentive programs; if the results suggest that incentives can have the desired impacts, then a more rigorous random assignment test of the concept could be contemplated.

schools before and after implementation of the incentive programs would be compared with changes over the same time period in a set of comparison schools selected to be as similar as possible to the high-need schools in terms of student demographic characteristics and prior academic achievement. It may be difficult to identify comparison schools that are closely matched with schools offering the incentives — since the closer the match, the more likely it is that the comparison schools would themselves offer incentives. Nevertheless, researchers should strive to find the best possible comparison schools and then employ regression controls in order to adjust for any systematic differences between the target schools and their comparisons, while acknowledging that it is not possible to control for unmeasured differences. The comparative interrupted time-series design, the selection of comparison schools, and several other design issues are considered in detail below.

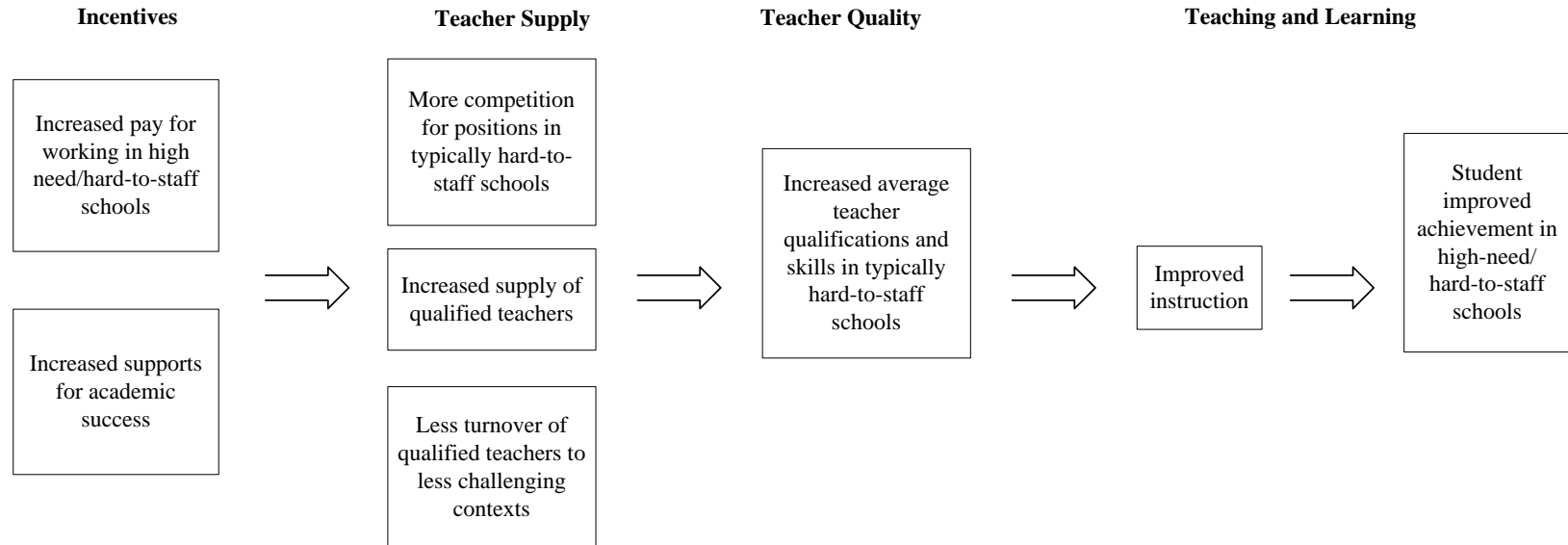
Before describing the study in depth, it is useful to explore the “theory of action” underlying teacher recruitment incentives, since this theory helps define the research questions that would drive the project. Figure 2 illustrates how recruitment incentives are hypothesized to affect teachers’ behavior. The basic logic is that, by compensating teachers for working in more difficult environments, school districts will experience two primary benefits. First, they will see an increase in the number of teachers willing to work in hard-to-staff schools. Second, they will see a reduction in the turnover that occurs when more experienced teachers leave to work in less challenging settings.

As the earlier review of various districts’ incentive programs indicates, incentives to work in high-need schools are not always targeted toward teachers with particular characteristics or qualifications. The theory implicit in these incentives is that a general increase in the supply of teachers to these schools (and the resulting competition for teaching slots) will result in an increase in the average “quality” — that is, skills and qualifications — of teachers working in these settings. The increase in supply is also hypothesized to lead to a reduction in the disparity between the qualifications of teachers working in historically low-performing schools (or schools serving high proportions of disadvantaged and minority students) and the qualifications of their counterparts in more advantaged environments. According to the theory, the result of having better-qualified teachers is, of course, improved teaching and improved student achievement outcomes, including better mastery of the particular skills assessed by locally administered standardized tests.

The remainder of this section focuses on the specific research questions the exploratory study would address, on recommended plans for data collection and analysis, and on how school districts might be selected to participate in the study.

Figure 2

Relationship Between Incentives and Teacher Quality



Primary Research Questions

The theory of action illustrated in Figure 2 suggests a set of research questions about context, implementation, and effects of staffing incentives on recruitment and retention outcomes in hard-to-staff schools.

Context

- What are the primary recruitment and hiring challenges faced by districts and schools considering or implementing teacher recruitment incentives?
- What is the existing (or “usual”) process for recruiting, hiring, and placing teachers, in the absence of any incentives to attract teachers to particular schools?
- To what extent do the average backgrounds and qualifications of teachers vary across different schools within these districts? To what extent do average qualifications vary from those of teachers in surrounding districts?
- How are teacher recruitment strategies affected by local circumstances, including the political climate, collective bargaining agreements, etc.?

Implementation

- What is the basic intervention design, and how does it change the pecuniary and non-pecuniary incentives for working in particular schools?
- What specific schools and teachers does the intervention target?
- How does the intervention fit into and interact with existing recruitment and hiring practices in the district and at the targeted schools?
- How was the intervention implemented, how were principals and teachers notified of the incentives, and what were the primary challenges, milestones, and successes in the implementation process?
- What are the “take-up” rates for the incentives — that is, how many teachers actually take advantage of the available incentives?

Effects on Recruitment and Retention Outcomes

- How does the “supply” of teachers applying to the targeted schools change in the years following implementation of the incentives?

- How do vacancy rates, average teacher qualifications, and length of retention change over time in the targeted schools?
- How do these changes differ from those occurring at the same time in similar schools (or schools that are as similar as possible) in the district? How do they differ from the changes in similar schools from nearby districts serving similar students?

Data Collection

In order to answer the research questions above, it will be important to go beyond quantified assessments of the net effect of the intervention(s) studied on recruitment and retention outcomes. Rather, the study should leverage several different types of data in order to flesh out the theory of action described above, to test hypotheses regarding the connections implied by this theory of action, and to provide useful information about the context, implementation, and effects of salary incentives on teacher recruitment and retention. In particular, the approach described below relies on field research, survey data, and administrative records in order to answer the key research questions.

Because this exploratory study focuses on intermediate outcomes (for instance, changes in teacher hiring and retention patterns) rather than the ultimate outcome of increased student achievement, the follow-up period for the study, and therefore for the data collection efforts discussed below, should include the first two to three years following the implementation of salary incentives in hard-to-staff schools.

Field Research

Field research — data collected from documents and from informed observers and participants (including those affected by the interventions) — can supply vital information about the context, origins and implementation of the reforms. For each intervention studied, the data to be gathered should concern:

- the political, organizational, and historical context in the sites where the incentive program is implemented;
- existing hiring and recruitment practices, key challenges, and important local circumstances;
- the origin and design of the initiative itself; and

- implementation of the initiative, including key milestones, perceptions of the effectiveness and of the incentive plans and the challenges they face, and the interaction between the new plans and the contexts in which they are put into place.

Document Review

Researchers should collect and review documents describing the existing recruitment and hiring practices, the intervention itself, and key facts about the local context. The review would include policy and strategy memos regarding hiring and recruitment, district strategic plans, and local press coverage of programs, events, and policies related to the school district. This review would be conducted early in the study to help inform and guide the rest of the field research effort.

Interviews

Researchers should conduct face-to-face interviews with an array of central office staff members, including the district leadership, those responsible for recruitment and hiring, and those involved in devising and/or implementing the incentive programs.

Focus Groups

Separate focus groups with teachers and principals in the participating districts can provide a better sense of how the incentives affect daily life in the schools. These focus groups would address the same broad issues as the interviews with central office staff but would provide the perspective of school-level actors on these issues.

Field research could be conducted over the course of two to three visits to each district. The first visit might focus on background and context, and the second visit on the nature of the incentive intervention and its early implementation. The third visit could trace the unfolding implementation story and follow up on questions raised by the first two visits. Ideally, field research activities would take place at comparison schools as well as schools where the incentives are in effect (although they might be somewhat more limited at the comparison schools). Field research data would be examined for consistency and frequency of themes in interviews and focus groups. These data would be used to answer the context and implementation questions posed above and to refine the execution and interpretation of the quantitative analysis.

Surveys

Teachers and principals at the target schools and at the comparison schools would be surveyed at several points over the course of the study. The surveys would cover some of the same topic areas as the field research but would allow for systematic collection of data from a much larger group of respondents. Thus, the surveys would address the following topics:

- What are the characteristics of teachers at a given school, in terms of experience, tenure, qualifications, and demographics?
- How are teachers recruited, hired, and placed at particular schools? What factors matter in the selection or placement process (i.e., tenure, experience, qualifications, choice)?
- What procedures allow teachers to move to other schools in the district? What factors matter (i.e., tenure, experience, qualifications, choice)?
- What types of formal incentives are in place to attract teachers to particular schools? What are teachers' perceptions of those incentives? How do they affect teachers' decisions about where and whether they teach?
- What incentives not currently in place would teachers respond to? Or what else about their jobs would need to change in order for the incentives to matter?

Administrative Records

This phase of research focuses on teacher recruitment and retention outcomes, suggesting an analytic emphasis on measuring the extent and nature of changes in two kinds of outcomes:

- Changes in the pool of teachers and applicants for teaching positions; and
- Changes in the qualifications of teachers working at hard-to-staff schools.

Information on these topics would come from administrative records available from the school districts being studied or from the states in which these districts are located.

Understanding how the teacher applicant pool changes over time is an important challenge for this research. In addition to exploring how teacher characteristics in hard-to-staff schools change, the research should also gather information regarding changes in the supply of teachers that are available and willing to work in these schools. The extent to which this is possible will depend on the specific teacher recruitment and assignment practices of the districts, as well as on the extent to which districts keep historical databases that reflect the hiring process. For example, if teachers apply for positions at particular schools within a district, the number and qualifications of those applicants may be recorded in some historical database. On the other hand, if teachers apply to the district as a whole (and the database reflects this fact), it will be more difficult to assess how the supply of high-quality teachers available to work in specific hard-to-staff schools changes over time. Where possible, data collection efforts should focus on such measures as the number of applicants per position, the number of vacancies remaining at the beginning of the school year, and the average qualifications of teachers applying to vacant positions.

The study should also directly address changes in the average qualifications and characteristics of teachers in typically hard-to-staff schools. Administrative records from school districts and state human resources and other systems should enable researchers to create measures of the following teacher characteristics:

- **Teacher experience** — while the research on the relationship between years of teaching and quality is mixed, virtually all compensation systems in public school districts around the country reward teachers with more years of experience.
- **Teacher education** — including college majors and possession of master’s degrees.
- **Teacher credentials** — including regular versus alternative certifications, subject matter certifications, and “advanced” credentials (e.g., certification by the National Board of Professional Teaching Standards).
- **Teacher “cognitive skill” and academic achievement** — including source of bachelor’s degree (since prestige of undergraduate institution has been used as a proxy for teacher achievement and skill).⁵⁸
- **Teacher effectiveness/“value-added”** — to the extent that teachers can be linked with particular schools and that student-teacher roster data are available on a historical basis, it may also be possible to examine the extent to which incentives are associated with changes in the effectiveness of teachers, as measured by each teacher’s performance history.⁵⁹

⁵⁸As discussed earlier, some authors have made the case that, more than other characteristics, what best predicts teaching excellence is the teacher’s general level of academic achievement. While some studies have used teachers’ college entry test scores as a proxy for academic achievement, these data are highly unlikely to be available for this study. However, the prestige of teachers’ undergraduate institutions has also been used in previous research as a proxy for this trait, and it has been shown to have a correlation with teacher effectiveness. Therefore, to the extent that information regarding teachers’ undergraduate institutions is available in state or district administrative records and can be connected to the schools at which teachers work, it can be used to create measures of the prestige of teachers’ undergraduate institutions. The study should then use these data to examine the extent to which the incentives are systematically linked to the presence of teachers with stronger academic backgrounds.

⁵⁹It should be noted that acquiring the data needed for such an analysis would be quite challenging, and the requisite data may not be available in many districts. Nevertheless, some states and districts maintain longitudinal databases that contain teacher-specific “value-added measures.” Moreover, it may be possible to use these databases to connect teachers to specific schools. Access to data such as these, beginning several years prior to the implementation of the incentive initiative, will be necessary in order to assess the relationship between the incentives and average changes in teacher quality as measured by value-added analysis.

- **Other measures of quality.** Most discussions of the disparities in teacher quality across more or less disadvantaged schools focus on objective characteristics such as education, credentials, and experience. However, some studies cited in the literature review of teacher quality presented in Section II refer to less objective and easily measurable characteristics, such as “contextual knowledge” and “perceived efficacy.” While the MDRC-Temple team would attempt to create survey measures of these constructs, it may not be feasible to assess the connection between incentives and these less objective dimensions of teacher quality.

Analysis Plan

A quantitative analysis of the relationship between incentive programs and teacher recruitment and retention outcomes should address two primary questions:

- How do teacher recruitment and retention outcomes (i.e., teacher supply and teacher characteristics) in hard-to-staff or high-need schools change over time after the implementation of incentives?
- How do these changes in recruitment and retention outcomes differ from those at similar schools from the same districts that do not have access to these incentives?

Together, the answers to these questions will provide valuable evidence about whether or not incentives targeted to particular schools are associated with meaningful changes in the ease of staffing high-need schools and in the qualifications of the teachers working there. The next section of the paper describes the comparative interrupted time-series design, which would be used to answer these questions. It is important to note that the proposed research design cannot completely isolate the effects of the incentives from other factors that could be driving changes in staffing outcomes at the high-need schools. If, however, the analysis reveals a systematic association between teacher recruitment and retention outcomes and the presence of incentives, the findings would justify a more rigorous study that would attempt both to isolate the effects of incentive initiatives on teacher characteristics and to measure the effects of the resulting changes in teacher characteristics on student achievement.

The Interrupted Time-Series Design

A comparative interrupted time-series analysis would involve two steps. First, changes over time in “teacher recruitment and retention outcomes” — indicators of teacher supply and characteristics — would be measured at schools where recruitment incentives are in place. Second, the analysis would compare changes at the “targeted” schools to the changes that occurred

in similar schools in the same (or a similar) district. The difference between these changes over time would provide an estimate of the effect of the financial incentives.

Measuring changes over time in teacher recruitment and retention outcomes at program schools

Ideally, one would like to measure the difference between teacher recruitment and retention outcomes that occur in schools that are targeted for recruitment incentives and the outcomes that would have been observed in these same schools had the incentives not been implemented. But since this is not possible, the primary analytic challenge is to develop a reliable estimate of the outcome levels that would have been observed in the absence of the intervention (i.e., a convincing “counterfactual”). If a true random assignment experiment cannot be put in place, the most effective way to approach this challenge may be to apply the logic of the comparative interrupted time-series (ITS) method.⁶⁰

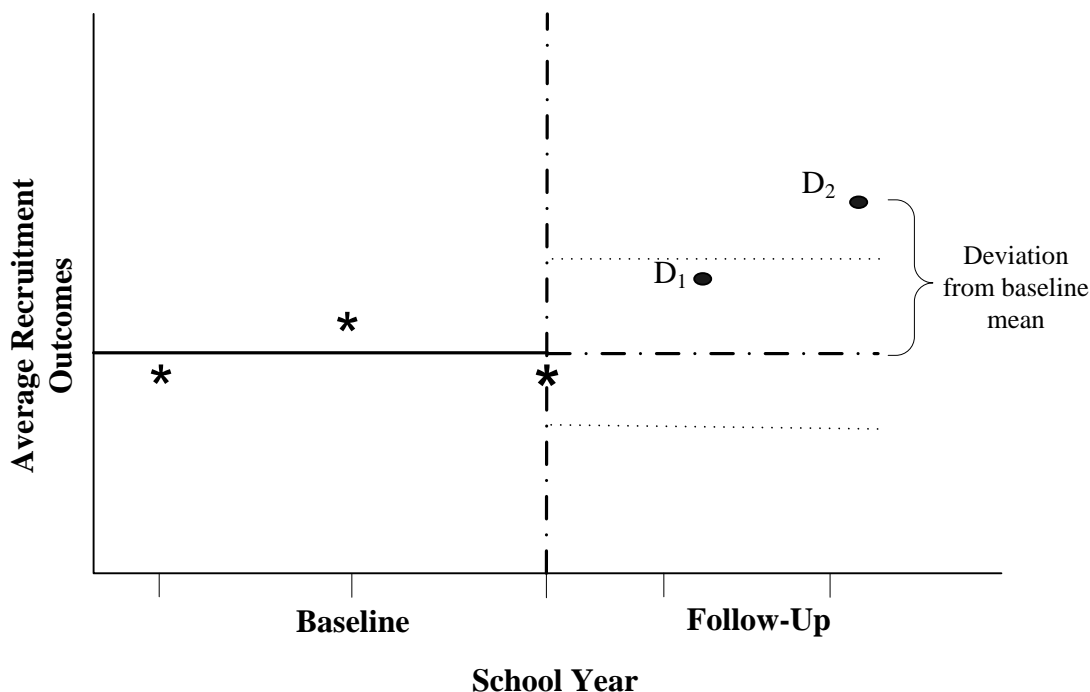
The fundamental logic of the ITS approach is that, absent any intervention or other major changes, the best predictor of future outcomes in a given school is the history of outcomes in that same school. The basic approach therefore rests on a comparison of the school-level changes over time in average recruitment and retention outcomes to baseline patterns in these same outcomes.

Figure 3, which illustrates this comparison, is a hypothetical example of the basic ITS method applied to a set of recruitment and retention outcomes in a fictional set of schools. The figure plots a teacher recruitment outcome — for instance, years of teaching experience — over a three-year baseline period and through two years after program implementation. The three asterisks on the left side of the graph represent the average value of this outcome across a set of “targeted” schools (that is, schools that are the focus of a district’s incentive program), during each of three years before implementation. The solid line running through these points represents the average value of the outcome over the entire three-year “baseline” period.

The right side of the figure illustrates recruitment outcomes during the follow-up period after the incentive program is put in place. The broken dotted line represents the value of the

⁶⁰See Bloom, H. (2003a). Using short interrupted time-series analysis to measure the impacts of whole school reforms. *Evaluation Review*, 27(1): 3-49.

Figure 3
An Illustration of Interrupted Time-Series Analysis



outcome that would be predicted based on the actual pre-program average. This projection serves as the benchmark against which post-program outcomes can be measured. The points on the graph marked “D1” and “D2” represent hypothetical average outcome levels in each year of the follow-up period. The distance between these points and the projected average — that is, the deviation from the baseline average — represents the first comparison in our analysis and answers the question, “Do teacher recruitment and retention outcomes change over time in schools that are the targets of incentive programs?”

The dotted lines on either side of the broken dotted line representing the projection from baseline establish a 95 percent “confidence interval” around the projected mean (i.e., the margin of error around the projection). To the extent that the average outcome value in any follow-up year falls outside the margin of error for this projection, one can be 95 percent certain that the deviation from the baseline average in that year is not merely due to chance. In short, this com-

parison tells us the extent to which teacher recruitment and retention outcomes have improved since the inception of the incentives.⁶¹

There are several important limitations on the inferences that can be drawn from this type of interrupted time-series analysis. For one thing, one cannot conclude with certainty that changes over time in average recruitment and retention outcomes in a particular set of schools were actually *caused* by the incentives. A variety of other factors may have driven these changes. For example, it is possible that, for reasons unrelated to the incentives, the composition of students changed over time at the schools targeted for the teacher incentives and that teachers were aware of these changes. It is also possible that a more “disadvantaged” or lower-performing group of students could make it harder to recruit qualified teachers, while more privileged students or students with better academic preparation would have the opposite effect. For that reason, it will be important for the analysis to control for average student characteristics.

The district may also intervene in these high-need or hard-to-staff schools in ways other than offering incentives for teachers to teach at these sites. For example, these schools might also begin to receive additional professional development or other instructional supports. These changes could make the schools more or less attractive to potential teachers and, as a result, could affect changes in the teacher recruitment and retention outcomes. Though this type of bias cannot be measured or controlled for, it will be important to use field research and survey data in order to ascertain the extent to which other developments at the schools affect the interpretation of the results. Moreover, to the extent that these changes are an explicit part of the program designed to change recruitment outcomes, this dynamic might not be thought of as “bias,” *per se*. Nevertheless, it would affect the interpretation of any estimated relationships between the incentive programs and teacher recruitment and retention.

Another key limitation relates to an issue sometimes referred to as “local history.” It is possible that changes that occur simultaneously with the implementation of incentives — for example, modification of state policy, the implementation of a new district-wide curriculum, or a change in district leadership — could exert a direct influence on recruitment and retention outcomes. Again, this makes it difficult to ascertain the extent to which any changes in these

⁶¹It is possible that, in addition to the average teacher recruitment outcomes, the manner in which these outcomes are changing over time, prior to any incentives, also vary across schools. As such, when possible, ITS analyses often control not just for the baseline averages but for the baseline trends — that is, rather than projecting future outcomes from a baseline mean, they project them from a baseline trend. However, such projections are not likely to be reliable without five years of historical data (see Bloom, 2003a). Therefore, to the extent that such historical data exists, analytic models should adjust to estimate baseline trends rather than baseline means. However, given that these data are unlikely to be available in all of the sites that would be included in this study, the research will probably have to rely on deviations from a baseline average. In this case, while researchers can control for observable differences in baseline averages, they cannot account for the fact that, even absent any intervention, these outcomes may have been changing over time.

outcomes are driven by the incentives or are the result of other local events and would have occurred anyway. In order to disentangle these phenomena, the design must extend beyond simple comparisons over time. This leads us to the second element of the comparative interrupted time-series approach.

Measuring differences in change over time in teacher recruitment and retention outcomes between program and comparison schools

Without a random assignment study, it is probably not possible to resolve all of the issues related to causal inference described above. However, in order to move towards distinguishing the changes in teacher recruitment outcomes that occur at schools receiving incentives from those that would have occurred regardless of the initiatives, the analysis should compare changes at the “targeted” schools to the changes that occurred in similar schools in the same district. In other words, for each follow-up year, the analysis should compare the average deviation from the baseline outcomes in the schools receiving the incentive intervention to the average deviation from the baseline average in a set of schools with similar performance histories and student demographic characteristics, drawn from the same district. Focusing on this difference rules out the possibility that estimates of the relationship between teacher recruitment and implementation of recruitment incentives are biased as a result of district-wide events that coincide with implementation of the incentives.

In order to execute this strategy, each “target school” should be matched with a set of comparison schools from the same district that are not eligible to receive any recruitment incentives. Ideally, comparison schools should be limited to those that are as similar as possible in terms of prior achievement, student demographics, size, and structure, but there may be several limitations on researchers’ ability to select closely matched comparison schools.⁶² First, incentives are typically targeted to schools with high populations of disadvantaged or minority students and/or a history of low performance. Most or all of the schools that are the best comparisons for the “target” schools may be eligible for the incentives themselves. If less suitable comparison schools must be chosen, there may be large and systematic differences between the student demographics and achievement levels at the targeted schools and the available comparisons. To the extent possible, the analysis will control statistically for some of these observable school characteristics, but it may not be possible to rule out the presence of systematic but unmeasured differences between the target schools and their within-district counterparts.⁶³

⁶²The study would also examine whether there are other factors about the comparison schools that make them fundamentally different from, and therefore unsuitable comparisons for, the schools in question (e.g., charter schools, schools on year-round schedules, etc.).

⁶³It is also worth noting that some districts implement district-wide incentives in order to attract high-quality teachers. Evaluations of these initiatives would have to rely on comparisons between schools in these
(continued)

While this analysis will not be able to establish with certainty whether incentives actually cause changes in teacher recruitment and retention outcomes, it will generate evidence about the magnitude of changes in teacher supply and teacher characteristics at schools targeted for incentives. It will also indicate whether or not these changes appear to be systematically different from the changes at schools from within the same context that are not the targets of incentives. Such evidence will be important in ascertaining whether large investments in demonstration studies of these incentives are justified and in refining the interventions and research design of such studies if they take place.⁶⁴

Sample Size and Minimum Detectable Effects

Minimum detectable effects (MDEs) are a simple way to express the statistical precision of an impact study design. Intuitively, a minimum detectable effect is the smallest program impact that could be measured with confidence given random sampling and estimation error.⁶⁵ For example, from a benefit-cost perspective, one might ask whether a proposed sample could reliably detect the smallest impact needed for a program to “break even” (that is, produce benefits equal to costs.) One would want a sample that was large enough to ensure that an estimated impact around this “break-even” point was a reliable indicator of the program’s true impact and not just due to chance variation. A smaller sample might only enable the study to detect impacts that are well above this break-even point and thus very difficult to attain. The study would, therefore, miss an opportunity to produce reliable estimates of policy-relevant impacts. Hence, it would be “under-powered” statistically.

With respect to statistical precision, the two most important issues for the analysis of incentives described in this paper are: (1) how to ascertain what kind of statistical power is

districts and similar schools in nearby (and hopefully similar) districts that do not offer the same salary incentives. In this case, while the analysis could rule out the possibility that differences in changes in recruitment and retention outcomes between targeted and comparison schools are due to state or regional context conditions, they could not rule out the possibility that these differences were driven by district-level factors unrelated to the incentives. For this reason, to the extent possible, research of this kind should focus on programs or initiatives that vary incentives within (rather than across) districts.

⁶⁴To the extent that districts assign incentives by ranking schools according to objective criteria and by picking a threshold that determines schools’ eligibility for incentives, it may also be possible to apply regression discontinuity designs in order to estimate the effects of the incentives on the outcomes of interest. See Thistlethwaite, D.L., and Campbell, D.T. (1960). “Regression discontinuity analysis: An alternative to ex post facto experiment.” *Journal of Educational Psychology*, 51(6), 309-17.

⁶⁵The MDRC-Temple team defines a minimum detectable effect as the smallest true program impact that would have an 80 percent chance of being detected using a two-tailed hypothesis test at the 0.10 level of statistical significance. The team uses a two-tailed test because the central policy issues to be addressed by this evaluation relate to whether specific interventions involving incentives have positive or negative effects on teacher recruitment outcomes.

needed, and (2) how statistical power varies with the number of schools included in an analysis sample.

The first issue essentially boils down to the question: “What is a policy-relevant change in teacher quality?” Ultimately, the most critical question about these incentives is whether or not they improve student achievement. The threshold for a policy-relevant effect on achievement is often considered to be about a 0.2 standard deviation.⁶⁶ If the incentives affect achievement, it will likely be through changes in the population of teachers at hard-to-staff schools. In order to have an effect on student achievement, these recruitment and retention outcomes (at least in the aggregate) will have to be at least moderately correlated with student achievement — say a minimum correlation of 0.5. This implies that, for any given effect on achievement, it would be necessary to generate an effect on teacher recruitment and retention outcomes that was twice as large. For example, in order to generate an effect of a 0.2 standard deviation on achievement it would be necessary to generate an effect of at least 0.4 on the teacher recruitment and retention outcomes that preceded achievement. This suggests that a target minimum detectable effect size for the teacher recruitment and retention outcomes described above should be 0.4.

With respect to the second question, estimates of minimum detectable effects for prospective studies are typically based on analysis of variance components of key outcomes in samples drawn from populations similar to the ones in the study at hand. For example, in studies focused on student achievement in urban school districts, researchers have drawn on samples from similar schools in similar school districts in order to estimate MDEs for their proposed studies.⁶⁷ In this case, the focus on teacher recruitment and retention outcomes may make this approach impractical. As the variance patterns for achievement outcomes may or may not be similar to the variance patterns for recruitment and retention outcomes (such as teacher experience, etc.), one cannot use the estimated MDEs from analyses of achievement in similar districts as an indicator of the MDEs yielded by a study of incentives on recruitment and retention in hard-to-staff schools. Moreover, data sets containing information on teacher characteristics and other recruitment and retention outcomes are less readily available for analysis. It is probably not possible to construct useful estimates of MDEs without this information. Attempts to collect these data and estimate MDEs based on several urban districts might be an important early stage in this research effort. In particular, data from existing studies of teacher recruitment and retention outcomes could be used to estimate the variation across schools in these outcomes and to generate estimated standard errors for an ITS study of the effects of the intervention.

⁶⁶Bloom, H. (2003b). *Sample Design for an Evaluation of the Reading First Program*. New York: MDRC.

⁶⁷See Bloom (2003b).

Selection of the Study Sites

Earlier sections of the paper described incentive programs in several districts around the country. In deciding which of these (or other) incentives programs to focus on, researchers should consider several factors, including the availability of longitudinal teacher recruitment and retention data, the nature and focus of the teacher incentives, and the magnitude of the incentives.

Availability of Longitudinal Teacher Recruitment and Retention Data

The analytic approach described in this paper rests on the ability to assess the extent to which changes in key teacher recruitment outcomes, such as the average years of teaching experience, average teacher education, and teacher certifications in each school, change in the years following implementation of incentives. The ability to carry out this analysis depends on the availability of databases containing information regarding teacher characteristics, the ability to link teachers to the particular schools in which they work, and the availability of data that precede the implementation of the reform itself. Previous research suggests that it may be possible to acquire these data from state-wide teacher information databases.⁶⁸ Though they may be difficult to obtain, having these data for at least three years prior to the reform in question is an important prerequisite to researchers' ability to study the relationship between incentives and changes in teacher characteristics in any particular district.

Nature/Focus of Incentives

Incentive programs vary in their approach as well as in their goals. Some focus on rewarding teachers for their individual performance in terms of the students they serve. Others reward all the teachers at a particular school for school-wide improvements in achievement. Still others reward teachers for participation in particular professional development activities, and some reward teachers for the completion of particular certification programs. Although there are many interesting questions to pursue about teacher incentives, this program of research focuses on the problem of differences in the supply of qualified and effective teachers across schools serving students with different demographic backgrounds and prior achievement — particularly the role of incentives (financial and otherwise) in addressing this problem through changes in the supply of qualified teachers interested in working in typically hard-to-staff schools. With that in mind, the analytic approach outlined above is most suited to districts in which specific financial (and perhaps non-pecuniary) rewards are targeted to teachers working in schools meeting particular criteria, or schools that are considered to be either “high need” or “hard to staff.”

⁶⁸See Cavalluzzo, L.C. (2004). *Is National Board Certification an Effective Signal of Teacher Quality?* Alexandria, VA: The CNA Corporation.

Magnitude of Incentives

What size of financial incentives should this study focus on? The reconnaissance presented in this paper underscores the fact that the magnitude of the incentives offered to teachers varies quite a bit depending on the district and the particular initiative in question. For example, while incentives to work in hard-to-staff schools are as large as 20 percent of a teacher's annual salary in Miami-Dade, the specific portion of the incentives associated with working in a high-need school is approximately three percent in the Denver Public Schools. Some might argue that smaller incentives are unlikely to create enough difference in teachers' willingness to work in particular schools to have a meaningful effect on average teacher characteristics and qualifications in hard-to-staff schools. At the same time, others might argue that large incentives are too expensive to be broadly applicable — and are therefore not policy relevant.

Selecting sites for a study is a balancing act among these tensions. In addition, researchers must decide whether they want to focus on a collection of districts with a range of incentives or on a subset of districts with particularly large (and potentially powerful) interventions. Teacher behavior is hard to change, and incentives are working in environments with other powerful forces, such as collective bargaining agreements and rules regarding the relationship between seniority and hiring and salary decisions. Moreover, averaging the effects of interventions that vary greatly in terms of their potential to affect teacher behavior seems problematic. Finally, there is not a lot of existing evidence suggesting that incentives can actually affect teacher recruitment and retention outcomes. With that in mind, researchers should consider focusing on the subset of districts that offer incentives that are perceived to be particularly powerful. At a bare minimum, in a study that includes several sites, the analysis should be done in a way that attempts to isolate the effects of incentives in each district. The variation across districts in the estimated relationships between the presence of incentives and changes in teacher recruitment and retention outcomes cannot be completely separated from idiosyncratic district-level events affecting teacher mobility and hiring. Nevertheless, describing the association between incentives and teacher recruitment and retention at districts that have made more substantial investments in attracting more qualified teachers to hard-to-staff schools would be an important step towards understanding the potential of these incentives to serve as a lever for improving achievement and equity in education.

Appendix

Potential Measures for Surveys of Teachers and Principals

- **Teacher background**
 - Years at school and in district
 - Total years teaching
 - Teaching credentials
 - College education, institution, and major
- **Degree of different types of challenges facing teacher at current school**
 - Students' low achievement
 - High mobility
 - Discipline problems
 - Lack of resources
 - Lack of parent and community support
 - Large class sizes
- **Reasons for teaching in current districts and at current school**
 - Salary
 - Resources for teachers (i.e., textbooks, supplies, planning time)
 - Achievement of students
 - Professional development/support
 - Status
 - Peers/peer support
 - Mentoring
 - Principal
- **Satisfaction with supports at current school**
 - Salary
 - Resources for teachers
 - Achievement of students
 - Professional development/support
 - Status
 - Peers/peer support
 - Mentoring
 - Principal

- **Ease or difficulty of moving to a more disadvantaged school or to a more advantaged school — based on meeting particular criteria**
 - Tenure
 - Experience
 - Credentials
 - Expertise in particular subject
- **Effects of district’s existing formal incentives (each separately and as a group) on individual teacher’s decision to remain at current school (for teachers currently in disadvantaged, hard-to-staff schools)**
 - Salary increment
 - More professional development, by type (i.e., associated with reading curriculum, using assessments, classroom management, etc.)
 - Better professional development, by type
 - Mentoring
 - Other
- **Effects of district’s existing formal incentives (each separately and as a group) on individual teacher’s decision to change to a low-performing school or to remain at current school (for teachers currently in less disadvantaged schools)**
 - Salary increment
 - Professional development
 - Mentoring
 - Other
- **If incentives do *not* matter to teachers, reasons that they do not matter**
- **What additional incentives would matter to teachers in order to remain at a disadvantaged school or move to a disadvantaged school?**
 - More salary
 - More professional development
 - Mentoring
- **Expectations about career future (i.e., in X years from current year)**
 - Expect to be teaching at same school
 - Expect to be teaching at different school in the same district
 - Expect to be teaching in another district
 - Expect to be an administrator or an educator in another capacity in same district
 - Expect to be an administrator or an educator in another capacity in a different district
 - Expect to leave teaching

About MDRC

MDRC is a nonprofit, nonpartisan social policy research organization dedicated to learning what works to improve the well-being of low-income people. Through its research and the active communication of its findings, MDRC seeks to enhance the effectiveness of social and education policies and programs.

Founded in 1974 and located in New York City and Oakland, California, MDRC is best known for mounting rigorous, large-scale, real-world tests of new and existing policies and programs. Its projects are a mix of demonstrations (field tests of promising new program approaches) and evaluations of ongoing government and community initiatives. MDRC's staff bring an unusual combination of research and organizational experience to their work, providing expertise on the latest in qualitative and quantitative methods and on program design, development, implementation, and management. MDRC seeks to learn not just whether a program is effective but also how and why the program's effects occur. In addition, it tries to place each project's findings in the broader context of related research — in order to build knowledge about what works across the social and education policy fields. MDRC's findings, lessons, and best practices are proactively shared with a broad audience in the policy and practitioner community as well as with the general public and the media.

Over the years, MDRC has brought its unique approach to an ever-growing range of policy areas and target populations. Once known primarily for evaluations of state welfare-to-work programs, today MDRC is also studying public school reforms, employment programs for ex-offenders and people with disabilities, and programs to help low-income students succeed in college. MDRC's projects are organized into five areas:

- Promoting Family Well-Being and Child Development
- Improving Public Education
- Promoting Successful Transitions to Adulthood
- Supporting Low-Wage Workers and Communities
- Overcoming Barriers to Employment

Working in almost every state, all of the nation's largest cities, and Canada and the United Kingdom, MDRC conducts its projects in partnership with national, state, and local governments, public school systems, community organizations, and numerous private philanthropies.