

NBER WORKING PAPER SERIES

WOULD HIGHER SALARIES KEEP TEACHERS IN HIGH-POVERTY
SCHOOLS? EVIDENCE FROM A POLICY INTERVENTION IN
NORTH CAROLINA

Charles Clotfelter
Elizabeth Glennie
Helen Ladd
Jacob Vigdor

Working Paper 12285
<http://www.nber.org/papers/w12285>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2006

We thank Kata Mihaly, Amber Gove and Audrey Beck for outstanding research assistance, Dale Ballou, David Figlio, and participants at the 2005 AEFA and APPAM meetings for helpful comments on previous versions, and the Spencer Foundation for financial support. Administrative data used in this study were collected by the North Carolina Department of Public Instruction and made available by the North Carolina Education Research Data Center. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2006 by Charles Clotfelter, Elizabeth Glennie, Helen Ladd and Jacob Vigdor. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Would Higher Salaries Keep Teachers in High-Poverty Schools? Evidence from a Policy Intervention in North Carolina

Charles Clotfelter, Elizabeth Glennie, Helen Ladd and Jacob Vigdor

NBER Working Paper No. 12285

June 2006

JEL No. I2, J33, J45

ABSTRACT

For a three-year time period beginning in 2001, North Carolina awarded an annual bonus of \$1,800 to certified math, science and special education teachers working in high poverty or academically failing public secondary schools. Using longitudinal data on teachers, we estimate hazard models that identify the impact of this differential pay by comparing turnover patterns before and after the program's implementation, across eligible and ineligible categories of teachers, and across eligible and barely-ineligible schools. Results suggest that this bonus payment was sufficient to reduce mean turnover rates of the targeted teachers by 12%. Experienced teachers exhibited the strongest response to the program. Finally, the effect of the program may have been at least partly undermined by the state's failure to fully educate teachers regarding the eligibility criteria. Our estimates most likely underpredict the potential outcome of a program of permanent salary differentials operating under complete information.

Charles T. Clotfelter
214A Terry Sanford Institute
Duke University
Box 90245
Durham, NC 27708
and NBER
charles.clotfelter@duke.edu

Elizabeth Glennie
North Carolina Education Research Data
Center
Duke University
Box 90545
Durham, NC 27708-0545
eglennie@duke.edu

Helen F. Ladd
214A Terry Sanford Institute
Duke University
Box 90245
Durham, NC 27708
helen.ladd@duke.edu

Jacob L. Vigdor
Terry Sanford Institute of Public Policy
Duke University
Box 90312
Durham, NC 27708
and NBER
jacob.vigdor@duke.edu

The twin topics of teacher quality and teacher compensation have garnered considerable attention from researchers and policy-makers in recent years. This attention has been motivated in part by the desire to increase the quality of individuals who select into the teaching profession, and to prevent attrition from culling the most qualified individuals from the ranks of teachers (Corcoran, Evans and Schwab 2004; Hoxby and Leigh 2004). A second motivation has been concern about the uneven distribution of effective teachers across schools. Numerous studies have documented the tendency for the most qualified teachers to gravitate toward schools that serve relatively well-off children, even though salaries are often no higher in such schools (Lankford, Loeb and Wyckoff 2002; Scafidi, Sjoquist and Stinebrickner 2002; Hanushek, Kain and Rivkin 2004; Reed, Rueben and Barbour 2006). As states and local districts feel increasing pressure to comply with the No Child Left Behind Act's mandate that each classroom contain a "highly qualified" teacher, evidence on policy interventions that successfully recruit outstanding candidates and distribute them equitably becomes more and more valuable.

One obvious policy tool to deal with the quality and distribution of teachers would be to increase teachers' monetary compensation, perhaps in a targeted way. As straightforward as that suggestion might seem, however, a large and growing body of evidence suggests that the power of higher salaries to attract better teachers to the profession is limited (Hanushek, Kain and Rivkin, 1999; Loeb and Page 2000), and that offering teachers pay differentials to take jobs in low-performing schools is not a cost-effective means of improving test scores, particularly when the pay differential consists of a one-time signing bonus rather than a permanent salary increase (Fowler, 2003). Signing bonuses, absent other contractual provisions, may influence the attraction of teachers to schools but offer no incentive for teachers to stay once they have arrived.

Whether it is because of this evidence, the opposition of teacher unions, or just the sheer expense, states and districts have shown little interest in *ongoing* financial bonuses or other forms of salary differentials designed to attract and keep qualified teachers in low-performing schools. This choice is somewhat surprising given the current policy interest in market-based reforms in education; these reforms usually involve the expansion of parental choice of schools and competition among schools. Yet the use of market incentives -- in the form of higher salaries or bonuses for teaching in underperforming schools -- could, in theory, generate large benefits to disadvantaged students. To be sure, the benefits of higher quality teachers come with the obvious costs of paying more money for teaching services, and the question of whether the benefits exceed the costs is inherently an empirical one. Moreover, the benefits of higher teaching quality are contingent on administrators' being able to pick the higher quality applicants from the enlarged pool, an ability that has been questioned by recent research (Ballou and Podgursky 1995).

To gain insight into these questions, this paper focuses on a program implemented in North Carolina from 2001/02 until 2003/04, which awarded annual bonuses of up to \$1,800 to certified teachers of math, science and special education in middle and high schools serving low-income or low-performing students (henceforth, either the North Carolina Bonus Program or, simply, the bonus program). The goal of the program was to make it easier for such schools to attract and retain qualified teachers in these fields, by funding a permanent within-district salary differential.¹ We provide a broad overview of the program and discuss results of surveys administered to eligible teachers and their superiors in Clotfelter et al. (2006).

¹ While the Bonus Program was intended to provide a permanent salary differential to teachers of eligible subjects in targeted schools, in practice the program was canceled after its third year of implementation.

Our analysis in this paper uses longitudinal data on North Carolina public schoolteachers to estimate discrete-time hazard models predicting teachers' decisions to end a spell of employment at a particular school. Turnover rates are an important outcome to consider because administrators often must fill vacancies in these schools with inexperienced teachers, who have been shown to be less effective in their first years on the job than otherwise similar teachers with more experience (Clotfelter Ladd and Vigdor 2006; Rivkin, Hanushek and Kain 2005; Rockoff 2004).² We identify the impact of the \$1,800 salary differential by incorporating data before and after the program's implementation, in schools that met and barely missed the eligibility criteria, and for teachers who taught targeted subjects and other subjects. The net result is a form of difference-in-difference analysis coupled with a regression discontinuity design. The results suggest that the sum of \$1,800 per year was sufficient to reduce turnover rates of the targeted teachers by roughly 12%. Moreover, communication failures led many teachers to underestimate the probability that they would receive a bonus payment if they remained at their school. Evidence suggests that these failures reduced the program's impact, making our result an underestimate of the potential change that would be brought about by a truly permanent salary differential.

Beyond considering the impact of more money, a question of equal or greater importance is whether the teachers influenced by such pay are the types that principals and other administrators would most like to retain. Our tests for heterogeneity in impact reveal that the bonus program was more effective in reducing turnover rates for teachers with ten to thirty years of experience. As experience is one of the few observable teacher characteristics that reliably

² Although the ultimate test of the Bonus Program would be whether it raised student achievement, our analysis of teacher retention rates suggests that any impacts would be too small to discern empirically except in samples significantly larger than those at our disposal. Hence, we focus here on an intermediate outcome, one which, had the program been maintained, might well have raised achievement further in the long run.

predict higher student achievement, this evidence further suggests that salary differentials may be an effective strategy for improving the quality of education in high-poverty schools. The bonus program also had stronger impacts on math teachers, and in middle schools rather than in high schools.

In section I of this paper, we provide a brief overview of the bonus program that encapsulates information contained in Clotfelter et al. (2006). In section II, we describe the data set we constructed for the survival analysis and the hazard function methodology. In section III, we present our basic results, both on the average effects of the program and on how those effects differ across various types of teachers and schools. In section IV, we analyze whether the implementation difficulties described in section I reduced the effectiveness of the bonus program. A brief conclusion in Section V completes the paper.

I. The Structure and Implementation of the North Carolina Bonus Program

The North Carolina Bonus Program took effect in September 2001. The state, which administers a centralized payroll system covering employees in each of the state's 117 districts, began providing an annual salary supplement of \$1,800 to teachers certified in math, science, and special education teaching those subjects in middle schools or high schools that met either of the following criteria:

- (1) 80 percent or more of students had to be eligible for free or reduced price lunch, or
- (2) 50 percent or more of its students had to perform below grade level in *both* Algebra 1 and Biology, as measured by the state's end-of course tests.

Part-time teachers and those who taught both targeted and non-targeted subjects received a prorated bonus amount. Uncertified teachers of math, science and special education received no

bonus payment. These teachers comprised roughly one-third of all teachers of these subjects statewide and a significantly higher fraction in some schools. Over the next three years, nearly two thousand teachers working in 148 schools in 65 districts were eligible for bonus payments in at least one year.

Although teachers needed to be employed in an eligible school to begin receiving the bonus, continued receipt did not require continued school eligibility. The only requirement was that the teacher continue to work in the same school, teaching one of the eligible subjects. This provision was intended to eliminate both uncertainty and any perverse incentives for teachers to keep test scores low so that the school would remain eligible.

For reasons described more fully in Clotfelter et al. (2006), the vast majority of teachers receiving bonus payments in the 2003/04 school year misunderstood the provisions of the bonus program, according to responses to surveys we administered in 2004.³ The most common misconception was that continued receipt of the bonus required that the school remain eligible. Thus, although the program was designed to assure teachers that they would continue to receive the bonus (if they did not change schools or subjects), many teachers did not understand this provision. Misperceiving teachers making rational decisions to stay or leave based on the expected stream of future bonus payments thus systematically erred, expecting payments lower than what the enabling legislation specified.⁴ For teachers who inaccurately believed that their continued receipt was linked to their school's continuing eligibility, the expected value of future

³ We administered surveys by mail, with telephone follow-up when necessary, to 1,165 principals and teachers in eligible schools. The response rate was 83% for principals and 72% for teachers. More detail on the survey and our findings can be found in Clotfelter et al. (2006).

⁴ Ironically, events proved the doubters correct. Despite the program's stated assurances, the state legislature canceled the bonus program after the 2003/04 school year. Thus skeptical teachers may also have – correctly – factored a lower probability of continued receipt into their calculations. The state's decision to cancel the bonus program was made rather abruptly in the summer of 2004. It is unlikely that the foreshadowing of this event directly impacted the teacher decisions analyzed in this paper, as the latest moving decisions factoring into our analysis would have occurred by Fall 2003.

bonus payments would be strongly linked to the criterion variable, either subsidized lunch receipt rates or algebra and biology test failure rates, determining that eligibility. Because the criterion variables fluctuate from year to year, a misperceiving teacher in a school with an 80.1% rate of subsidized lunch receipt, for example, would have had a lower expected value of future bonus payments than a similarly misperceiving teacher in a school with a 90% rate. In Section IV below, we will use this fact to test whether misperceptions altered the impact of the bonus program.

Straightforward economic theory suggests that the extra pay provided through this bonus program should increase the supply of teachers willing to work in an eligible school, and by extension reduce the rate of departure for teachers currently employed at that school. The magnitude of the impact is inherently an empirical question. The less elastic the labor supply curve for teachers in disadvantaged schools, the less an effect we should expect to observe.

II. Data and Methods

To estimate how the bonus affected the decisions of teachers about whether to remain in particular schools, we estimated a series of hazard models. Such models predict the probability that a teacher will depart from a school in year $t+1$, conditional on her being in the school in year t . The key explanatory variable in each model is a dichotomous variable indicating whether or not the teacher was paid a bonus in the particular year. Most of the estimated models include a variety of teacher and school characteristics as control variables. In addition, in some models, we interact the bonus indicator variable with selected teacher or school characteristics to test for differences in program impacts.

Our identification strategy has three basic components: we compare hazard rates (1) before and after the implementation of the bonus program, (2) across eligible and ineligible teachers working in the same schools, and (3) across teachers in eligible schools and in schools that narrowly missed the eligibility criteria. Ours is thus a difference-in-difference-in-difference strategy, with the third difference resembling a hybrid between a randomized experiment and a regression discontinuity design. We now describe in some detail the process of designating “treatment” and “control” schools.

We ranked schools by averaging the criterion variables (percent of students receiving subsidized lunch for middle schools and percent of test-takers failing Algebra I and Biology exams for high schools) across four years, the two used to determine eligibility in the first two years of the bonus program and the two previous years, 1998/99 through 2001/02. Because the criterion variables fluctuate from year to year, some schools appearing in the “treatment” category have averaged criterion variable values below the eligibility threshold. For example, a middle school could have a long-run average subsidized lunch rate of 75%, yet exhibited a rate of 81% in a single year, which would have triggered bonus eligibility in the subsequent year.⁵ Since there is not a strict threshold dividing eligible from ineligible schools using this metric, ours is not a true regression discontinuity design. The departure is in the direction of a true randomized experiment, however.

To form a set of control schools, we selected a set of middle and high schools with mean criterion variable values comparable to those of treatment schools. The results of the selection process are summarized in Table 1. Treated middle schools had average criterion variables –

⁵ Interestingly, the “treatment” schools in our analysis tend to be smaller than the “control” schools, which is to be expected when eligibility is based on a noisy indicator. Smaller schools are more likely to have criterion variable values that depart markedly from the mean in any particular year.

subsidized lunch receipt rates – that ranged from 0.633 to 0.932. Recall that 0.8 is the eligibility threshold, based on a single years’ data point. Our selected group of control middle schools had average criterion variables in the range from 0.725 to 0.841.⁶ While there is a noticeable difference between the means of the average criterion variables in the treatment and control samples, the distributions overlap completely. We excluded middle schools with average criterion variables below 0.725, which is a natural break point in the distribution. We also excluded middle schools that came into existence or ceased to exist at some point during the sample period. Average criterion variables tended to be much lower in these excluded schools.

Among eligible high schools, failure rates on biology and algebra test scores tend to be quite a bit higher than the 50% eligibility threshold, particularly for biology exams. Note, however, that there are eligible schools where the failure rate averaged over several years is below the threshold. Selected control schools have average failure rates that generally fall close to the eligibility threshold. Because high schools had to have failure rates in excess of 50% on both exams to be eligible, it was possible to have very high failure rates on one exam without attaining eligibility. Summing the average failure rates of the two exams to form a criterion variable, each of the control high schools has a value of 83% or above.⁷ There is a natural break in the distribution of criterion values just below this level. Treatment schools have criterion variable values as low as 77%. High schools excluded from the analysis tend to have average failure rates well below the eligibility threshold.

⁶ Control group schools can have means that exceed the eligibility threshold if the criterion variable took on its highest values in the years before the bonus program was initiated. Also, some schools included in the control group became eligible for the bonus program in 2003/04, a year not included in our analysis. Excluding these schools from the control group produces no significant change in the quantitative results.

⁷ This information is not directly recoverable from Table 1, as the control school with the lowest biology failure rate was not the same as the school with the lowest algebra failure rate.

Finally, note that our treatment and control groups as constructed consist primarily of high schools, even though the majority of secondary schools in the state of North Carolina were middle schools. This likely indicates that the eligibility criteria were constructed in such a way that made it easier for a high school to qualify than a middle school.

Our longitudinal dataset on teacher workplace locations begins with the 1999/2000 school year, two years before the bonus program was implemented. Since the bonus program was not passed into law until the 2001/02 school year had already started, we will consider teacher decisions made following the first two years of our panel to be uninfluenced by the program. The latest year for which administrative data is currently available is 2003/04. Hence we constructed a four-year panel dataset, where the outcome of interest is whether a teacher employed in a particular school in year t remains employed at the same school in year $t+1$.⁸ Summary statistics for the teacher/year variables included in the analysis appear in Table 2. The teachers in our sample are roughly 65 percent female, 34 percent black, and 4 percent Hispanic or other nonwhite. These racial proportions differ in important respects from the overall population of secondary school teachers in North Carolina; in the excluded schools, only about 12 percent of the teacher/year observations are for black teachers.

Within the treatment and control schools, about seven percent of the teacher/year observations represent individuals in their first year of the profession, the median teacher has

⁸ In our data, we observe multiple instances of teachers leaving a school in year $t+1$ and returning to the same school in year $t+2$. These temporary absences most likely reflect maternity leaves or analogous situations. We treat any departure from school, temporary or not, equivalently in our analysis. The clearest justification for adopting this rule is our lack of knowledge regarding the ultimate outcomes of teachers who leave a school after the 2002/03 school year. Some of these teachers may return to the same schools in 2004/05, a year for which we do not currently have data. Adopting our convention is the only way to ensure that all departures are treated consistently regardless of the year in which they occur. When a teacher returns to work in a school, or begins work in a different school, we consider it the beginning of a new “spell” for purposes of our hazard model.

between 10 and 19 years of experience, and 29 percent of the teachers in these schools had some form of graduate degree. Roughly 9 percent of our observations correspond to teachers who received a bonus payment in the relevant year.

We estimate Cox proportional hazard models.⁹ Defining the conditional probability of leaving a teaching position at the end of year t as λ_t , the Cox proportional hazard model predicts the logarithm of the odds ratio, $\ln(\lambda_t/(1-\lambda_t))$, as a linear function of independent variables. In the tables below, we report the estimated coefficients as hazard ratios. Ratios above one correspond to variables that positively covary with the probability of departure, and ratios below one indicate variables that covary negatively with the probability of departure. Our primary goal is to test the hypothesis that, all else held constant, teachers who receive the bonus are less likely than other teachers to depart from their position in a particular year. Thus, the test is whether the hazard ratio is less than one for such teachers.

We rely on administrative payroll records to identify which teachers received the bonus each year. In principle, the payroll records should correspond to eligibility as determined by the requirements of the program. As described above, these requirements are that the teacher be certified in and teaching a targeted subject at least part time in schools that currently meet the eligibility criteria, or have been certified and teaching a targeted subject continuously in a school since a time when it was eligible. Using information on teacher licensure, employment, and teaching assignments, we have attempted to replicate the payroll records with limited success: 15 percent of teachers who appear to meet the eligibility criteria are not recorded as receiving the bonus, and 25 percent of teachers receiving the bonus do not appear to meet the eligibility

⁹ In principle, it would be interesting to estimate a second model analyzing whether the prospect of receiving a bonus program encouraged teachers to take on a position in an eligible school. As discussed in Clotfelter et al. (2006), the bonus program's structure made it a poor tool for the recruitment of teachers. School eligibility for a given academic year was usually not announced until that year had already started. For this reason, many school and district administrators did not advertise the bonus program in their efforts to recruit teachers.

criteria. It is unclear at this time whether these discrepancies reflect errors in our coding of eligible teachers, or errors in the administration of the program itself. Although we estimated models using both approaches to identify who was eligible for a bonus, we report results mainly based on the official payroll records.

Under some circumstances, such as teaching an eligible subject part time, some teachers received prorated bonuses. In some alternative models, we replaced the indicator for whether the teacher received any bonus with the amount of the bonus received. Because the decision to teach an eligible subject part time may be correlated with other unmeasured individual characteristics, however, we relegate those results to footnotes.

III. Results

Simple difference-in-difference results

Table 3 presents a basic hazard model controlling for three variables: a binary indicator for whether a teacher received a bonus payment in a particular year, a second indicator for whether the teacher was certified in one of the subjects targeted by the bonus program, and a binary control for whether a school met the eligibility criteria for the bonus program in the year in question. The hazard modeled in this specification, as well as later specifications, is departure from the school of current employment after the end of the year.

In this simple specification, indicator for bonus receipt is effectively a two-way interaction of the other two binary indicators, implying a difference-in-difference identification strategy. The eligible school main effect is significantly greater than one, indicating relatively high turnover rates in those schools. This is one indication that the “control” schools in our

analysis were not fully comparable to the treatment schools. The individual eligibility indicator, which applies to certified teachers in both eligible and ineligible schools, is also greater than one, albeit not significantly. The interaction term, which compares the relative turnover of eligible versus ineligible teachers in eligible versus ineligible schools, is statistically significantly less than one at the 10% level. Relative to other teachers in the same schools, teachers who received bonus payments were ten percent less likely to depart at the end of the school year.¹⁰

More refined estimates of the average program impact

The specifications reported in Table 4 move beyond the simple partial correlations and difference-in-difference style estimates by incorporating a growing number of additional covariates. To begin, we introduce controls for the subject(s) taught by each teacher in each year.¹¹ Although coefficients associated with these controls are not included in the table, results show significant variation across subject areas. Math, special education, and foreign language teachers are significantly more likely to depart, while history, practical arts, and technical/vocational teachers are significantly less likely to leave. Controlling for subject area leads to a modest increase in the estimated impact of the bonus program: recipients are now estimated to be 13 percent less likely to depart, other things equal. The hazard ratio is now significantly less than one at the 5 percent level.

In the second column of Table 4, we introduce controls for teacher gender and race. While there is no significant difference in the departure rates of male and female teachers,

¹⁰ Comparable difference-in-difference estimates using alternative bonus measures yield somewhat different results. Using our proxy measure of whether a teacher should have been eligible for the bonus, rather than payroll records indicating whether a teacher actually received a bonus, we find a hazard ratio insignificantly different from one. Using the dollar amount of the bonus in place of a binary indicator for receipt, we find a statistically significant pattern suggesting that teachers receiving the full \$1,800 were 15% less likely to depart, while those receiving a prorated bonus of \$900 were 8% less likely to depart.

¹¹ Note that these variables are not mutually exclusive since teachers can teach multiple subjects in the same year. Hazard ratios associated with these variables should be interpreted as the marginal impact of teaching a particular subject in addition to the set of subjects one already teaches.

Hispanic and multiracial teachers display higher hazard rates, and Native American teachers lower rates. Once again, these controls lead to a very modest change in the estimated impact of the bonus program, with no change in its statistical significance.

Controls for teacher experience and education levels, introduced in the third column, reveal patterns consistent with many existing findings on teacher departure rates. Teachers in their first year in the profession are 52 percent more likely to depart than those with 10 to 19 years of experience; departure rates are somewhat lower but still elevated for teachers in their second or third years. Departure rates rise significantly once teachers reach 30 years of experience. Teachers with graduate degrees are significantly more likely to depart schools in this sample, perhaps because these teachers have access to a wider array of opportunities in other schools and districts. Controlling for these factors has very little effect on our estimate of the impact of the bonus program.

The final specification permits hazard rates to differ for middle and high schools. A substantial difference emerges across the two school types. The average likelihood of departing in any year is 12 percent lower for teachers in high schools than for those in middle schools, other things equal. Controlling for this factor reduces our estimate of the program impact slightly, to a 12 percent reduction in the probability of departure, conditional on remaining in the school until the year in question. This hazard ratio of 0.88, which is significantly different from one at the five percent level, represents our best estimate of the overall mean impact of the bonus program on the retention of eligible teachers.¹²

¹² Estimates employing our alternative proxy measure of teacher eligibility, rather than official records indicating bonus receipt, continue to be statistically indistinguishable from one. Estimates using the dollar value of the bonus, rather than a binary measure of receipt, imply that the effect of receiving the full \$1,800 is a 17% reduction in the hazard rate, while the impact of a prorated \$900 reduction would be 9%. This in turn suggests a rough rule of thumb that a \$100 increase in the bonus reduces the probability of departure by approximately 1% -- not one percentage point, but one percent.

Variation in program impacts

Although the mean impact of the program is an important measure of program impact, it is also helpful to know whether the program's effects were selective in the sense of decreasing the likelihood of departure more for some types of teachers than for others. We focus our attention here on the potential differential impacts of the bonus program among teachers who differ along easily observable dimensions, as well as differential impacts by year of implementation and by type of school. In principle, it would also be desirable to examine differences by teacher effectiveness. Although some studies have estimated measures of teacher effectiveness in the form of so-called teacher fixed effects, in practice, that approach to estimating the effectiveness of individual teachers is not feasible for this study.¹³

We estimate the differential effects by interacting the various characteristics of interest with whether the teacher received a bonus in the particular year. The results are shown in Table 5. The first specification in that table considers whether the program had a stronger impact in its first or second year of implementation. The results suggest that the impact of bonus receipt on the conditional probability of remaining in the school was strongest in the first year of implementation, 2001/02. Teachers receiving the bonus in that year were 17 percent more likely to remain in the same school in 2002/03, relative to ineligible teachers or teachers in ineligible schools.¹⁴ Point estimates suggest an effect of roughly half the magnitude in the following year. The estimates of the program's impact in the second year blends the impact on teachers who

¹³ In North Carolina, standardized tests in high schools are generally administered at the end of specific courses, rather than to all students at the end of a grade. Thus, only a subset of teachers can be associated with student test scores, and even for these teachers there is a sample selection concern, as less academically proficient students will often not enroll in courses with a testing requirement. While more general end-of-grade tests are administered in middle school, in most cases we lack the ability to match students to teachers. Moreover, standardized tests in math and reading will provide little or no information regarding the quality of instruction in science and other subjects.

¹⁴ Estimates from an alternative specification relating the amount of the bonus received to the conditional probability of departure suggest that teachers receiving the full \$1,800 bonus in the program's first year were 25% less likely to leave by 2002/03.

were receiving the bonus for the second time with newly eligible teachers. One possible explanation for the reduced effectiveness may be that the quality of information available to newly eligible teachers was poor relative to what teachers eligible the first year were told. It should be noted, however, that it is not possible to reject the hypothesis that the program's impact was the same in both years.

We next examine how the response to the program differed among the eligible teachers depending on the subject they teach. A clear significant pattern of selective impacts appears. Math teachers receiving the bonus were 18 percent less likely to depart the following year when they received the bonus.¹⁵ Point estimates suggest no response among teachers of science or special education. This finding is particularly interesting in conjunction with the higher exit probability for math teachers noted in Table 4. It thus appears that the bonus program was particularly effective in lowering the hazard rates for those teachers at highest risk of departure.¹⁶

The third column in Table 5 presents an intriguing pattern of differences in the program impact by the experience level of the teacher. Point estimates suggest that responses to the program were concentrated among more experienced teachers, with a statistically significant 37 percent reduction in departure rates for teachers with 10 to 19 years of experience.¹⁷ Teachers with fewer than 10 years of experience appear to have been less influenced by the bonus payment. It should be noted, however, that we are unable to reject the hypothesis that the program impact was the same across all experience categories. Nonetheless, given previous findings indicating greater classroom effectiveness of experienced teachers relative to novice

¹⁵ Estimates from an alternative specification relating the amount of the bonus received to the conditional probability of departure suggest that math teachers receiving the full \$1,800 bonus were 24% less likely to leave.

¹⁶ An alternative explanation is that math teachers perceived the eligibility criteria for the bonus program more accurately, and thus considered it to be more valuable than their counterparts in other subjects.

¹⁷ Estimates from an alternative specification relating the amount of the bonus received to the conditional probability of departure suggest that teachers with 10 to 19 years of experience receiving the full \$1,800 bonus were 43% less likely to leave.

teachers (Clotfelter, Ladd and Vigdor 2005b; Rivkin, Hanushek and Kain 2005), and the greater propensity for novice teachers to serve disadvantaged students (Clotfelter, Ladd and Vigdor 2005a), this pattern shows the importance of further investigation into the potential role of monetary rewards as a tool for retaining experienced teachers in disadvantaged schools.¹⁸

The final set of interaction results sheds light on the differential impact of the bonus program by type of school.¹⁹ A large and statistically significant difference emerges. Middle school teachers receiving a bonus payment were 27 percent less likely to leave at the end of the year, while high school teachers exhibited no statistically significant response to the program.²⁰ Once again, this result is interesting in light of the observation in Table 4 that teachers in middle schools had significantly higher average likelihoods of departure than did teachers in high schools.

Taken at face value, these results suggest that the cost-effectiveness of salary differentials could be increased by targeting them towards relatively experienced math teachers in middle schools. The possibility remains, however, that variation in responses to the bonus program reflected variations in teachers' perceptions of the program's generosity, rather than differences in true behavioral parameters. The following section investigates the potential role of misperceptions in undermining the bonus program's impact.

¹⁸ An alternative explanation for this pattern is that experienced teachers may have been more likely to understand the eligibility criteria, either because of their greater experience with public school bureaucracy or because they tended to teach in schools where the information flow was more reliable. Our survey evidence suggests that experienced teachers were more likely to be aware of the program (Clotfelter et al. 2006).

¹⁹ In additional unreported specifications, we tested for differential impacts of the bonus program by teacher race, gender, and education level. Point estimates suggest that the program had its largest impact on white teachers, female teachers, and teachers with masters degrees. In no case are we able to reject the null hypothesis of equal impact across categories.

²⁰ Estimates from an alternative specification relating the amount of the bonus received to the conditional probability of departure suggest that middle school teachers receiving the full \$1,800 bonus were 31% less likely to leave.

IV. The Potential Role of Information

Following the discussion in Section I above, we conducted additional hazard models that introduced the bonus eligibility criterion variable – the same factor used to select control group schools in our analysis – and interacted that variable with the teacher-level indicator of bonus eligibility. Eligible schools with lower levels of the criterion variable were more likely to revert to ineligible status in a subsequent year. Teachers who fully understood the structure of the bonus program would have realized that this was of no concern to them. If all teachers had understood the operation of the bonus program, therefore, there would be no reason to expect the program to have been less effective at marginally eligible schools. If we consider working conditions in a school to vary inversely with the criterion variable, then it actually would be reasonable to expect a larger impact in marginally eligible schools. For example, suppose working conditions in a marginally eligible school can be described as a scalar value a , and in unambiguously eligible schools as a value $b < a$. Given an equal bonus amount in the two schools, it is reasonable to think that a larger number of teachers would consider the bonus adequate compensation for accepting conditions a rather than b .

If, on the other hand, teachers widely misperceived the eligibility criteria, then the program's impact on retention rates might well be lower at marginally eligible schools. If this were the case, estimates of the program's impact based on schools that were markedly above the eligibility threshold might be better indicators of what a fully comprehended bonus program might accomplish.

Column (1) in Table 6 presents a hazard model that introduces the criterion variable alone. The criterion variable has been normalized so that a value of zero corresponds to a school where the long-run average value equals the eligibility threshold. For middle schools, this is a

subsidized lunch receipt rate of 85%. For high schools, it is a combined failure rate on biology and algebra exams of 100% (on a scale from 0 to 200%). This is a relatively crude method of normalization, but provides a relatively straightforward interpretation. Note that including a linear control for the criterion variable makes our estimated equation more like a true regression discontinuity design.

In this specification, the estimated mean impact of the bonus program is very similar to previous estimates, a 14% reduction in turnover rates. The criterion variable enters significantly and positively, consistent with the presumption that higher values of the criterion variable correspond with less favorable working conditions at schools. The point estimate indicates that middle schools with 90% subsidized lunch rate exhibits turnover rates 3 percent higher than schools with 85% subsidized lunch rates, or that high schools where the failure rates on algebra and biology exams are 55% each have turnover rates 6% higher than in schools where the failure rates are 50% each.

In column (2), we introduce an interaction between bonus receipt and the criterion variable. The main effect of bonus receipt, which now represents the predicted impact in schools where the average criterion variable equals the eligibility threshold, is a statistically insignificant 10 percent. The interacted effect is estimated imprecisely, with a point estimate clearly but insignificantly below one. Taken at face value, this point estimate implies that the impact of the bonus program on turnover rates increased by one percent for every five percentage point increase in the criterion variable. Given the range of the criterion variable among middle schools is roughly 0.3, this implies that the bonus program was more than twice as effective in the eligible school with the highest criterion variable than in the school with the lowest criterion

variable. The range among eligible high schools was even greater, suggesting an even stronger differential impact.

How large might the bonus program's impact have been if all teachers had correctly perceived the eligibility criteria? Without formally modeling the process of teacher's expectation formation, any estimate we provide would be little more than an educated guess, particularly given the imprecision of the estimates in Table 6. But it is not unreasonable to think that the reduction in turnover rates could have been twice the magnitude in a perfectly informed world. In sum, this analysis provides suggestive, but not conclusive, evidence that the state's failure to fully inform teachers of the program's eligibility criteria at least partly undermined its effectiveness. Thus the estimates we present of the program's impact most likely understate the potential impact of a program operating under conditions of complete information.

V. Concluding Discussion

The ultimate goal of the N.C. Bonus Program was to improve the quality of math, science, and special education instruction for students in disadvantaged schools. The fact that the program appears to have reduced departure rates of teachers from the schools serving disadvantaged and low-performing students means that the program could potentially have raised student achievement had it remained in operation for a longer period of time. To the extent that higher teacher turnover rates disrupt the delivery of educational services within a school, lower turnover rates by themselves would be beneficial to students. In addition, lower turnover rates are likely to reduce the hiring of novice teachers. Given that novice teachers are generally less effective than more experienced teachers (Clotfelter, Ladd and Vigdor 2005b), any reduced reliance on novice teachers in such schools would provide an additional benefit to students.

Unfortunately, however, the North Carolina Legislature ended the program even before the first evaluation of the program was complete. Hence, it will not be possible to use the North Carolina experience to determine the full potential of such a bonus program.

Despite its premature demise, the North Carolina Bonus Program provides an important example of a type of reform that deserves more attention. Though we were able to examine only one outcome of the program in this study, the fact that the program generated positive effects, especially for math teachers in middle schools -- despite the flaws in its implementation which quite possibly diluted its impact -- bodes well for programs designed to target permanent salary differentials to certain types of teachers in needy schools.

REFERENCES

Ballou, Dale and Michael Podgursky, "Recruiting Smarter Teachers," *The Journal of Human Resources* 30 (Spring 1995), 326-338.

Clotfelter, Chalres T., Elizabeth Glennie, Helen F. Ladd and Jacob L. Vigdor. 2006. "Teacher Bonuses and Teacher Retention in Low Performing Schools: Evidence from the North Carolina \$1,800 Teacher Bonus Program." Unpublished paper.

Clotfelter, Charles T., Helen F. Ladd and Jacob L. Vigdor. 2005a. "Who Teaches Whom? Race and the Distribution of Novice Teachers." *Economics of Education Review* v.24 pp.377-392.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2005b. "Teacher Sorting, Teacher Shopping, and the Assessment of Teacher Effectiveness," *Journal of Human Resources*, forthcoming.

Corcoran, Sean P., William N. Evans, and Robert M. Schwab, "Changing Labor-Market Opportunities for Women and the Quality of Teachers, 1957-2000," *American Economic Review Papers and Proceedings* (May 2004), 230-235.

Fowler, R. Clarke. 2003. "The Massachusetts Signing Bonus Program for New Teachers: A Model of Teacher Preparation Worth Copying?" *Education Policy Analysis Archives* 11 (April 22). Retrieved 10/4/04 from <http://epaa.asu.edu/epaa/v11n13/>.

Rivkin, Steven G., Eric A. Hanushek and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement," *Econometrica* 79 (March) pp.417-458.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2004. "Why Public Schools Lose Teachers," *Journal of Human Resources* 39 (Spring 2004), 326-354.

Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 1999. "Do Higher Salaries Buy Better Teachers?" NBER Working Paper #7082.

Hoxby, Caroline M. And Andrew Leigh, "Pulled Away or Pushed Out? Explaining the Decline of Teacher Aptitude in the United States," *AER* 94 (May 2004), 236-240.

Ladd, Helen F. 2002. *Market Based Reforms in Education*. Washington, D.C.: Economic Policy Institute.

Lankford, Hamilton, Susanna Loeb, and James Wyckoff, "Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis," *Educational Evaluation and Policy Analysis* 24 (Spring 2002), 37-62.

Loeb, Susanna and Marianne E. Page. 2000. "Examining the Link Between Teacher Wages and

Student Outcomes: The Importance of Alternative Labor Market Opportunities and Non-Pecuniary Variation,” *The Review of Economics and Statistics* 82 (August), 393-408.

Murnane, R.J., J.D. Singer, J.B. Willett, J.J. Kemple, and R.J. Olsen. 1991. *Who Will Teach? Policies that Matter*. Cambridge: Harvard University Press.

Olson, Lynn. 2000. “Sweetening the Pot,” *Education Week*. Retrieved 10/4/04 from <http://www.edweek.com/sreports/qc00/templates/article.cfm?slug=recruit.htm>.

Prince, Cynthia D. 2002. “Higher Pay in Hard-to-Staff Schools: The Case for Financial Incentives,” unpublished paper, American Association of School Administrators, Arlington, VA (June). Retrieved 10/4/04 from http://www.aasa.org/issues_and_insights/issues_dept/higher_pay.pdf.

Reed, Deborah, Kim S. Rueben, and Elisa Barbour, *Retention of New Teachers in California* (X: Public Policy Institute of California?, 2006) Retrieved 4/14/06 from www.ppic.org/content/pubs/R_206DRR.pdf,

Rockoff, Jonah E., “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data,” *American Economic Review Papers and Proceedings* (May 2004), 247-252.

Scafidi, Benjamin, David Sjoquist, and Todd R. Stinebrickner, “The Impact of Wages and School Characteristics on Teacher Mobility and Retention,” Unpublished paper, September 2002.

Table 1: School Level Summary Statistics

	Treatment Schools	Control Schools	Excluded Schools
Proportion High Schools	0.631 (n=103)	0.625 (n=56)	0.382 (n=791)
Proportion Eligible for Free or Reduced Price Lunch (middle schools)	0.821 [.633 , .932] (n=38)	0.767 [.725 , .841] (n=21)	0.421 [0 , .995] (n=488)
Failure rate on Biology Exam (high schools)	0.743 [.479 , 1.000] (n=65)	0.501 [.310 , .938] (n=35)	0.348 [0 , 1] (n=284)
Failure Rate on Algebra Exam (high schools)	0.648 [.233 , 1.000] (n=65)	0.474 [.273 , .708] (n=35)	0.238 [0 , 1] (n=284)

Note: Unit of observation is the school. A school is defined to be in the “treatment” group if it met the eligibility criteria for inclusion in the bonus program in the first or second year of implementation. “Control” group schools were not eligible in either of the first two years, but approached the thresholds for eligibility. Five of the 56 control group schools became eligible for the bonus program in its third year. The range is in brackets, and the number of observations are in parentheses

Table 2: Teacher-Year Level Summary Statistics

	Mean
Subject Taught (no omitted category):	
Math	0.225
Science	0.184
English	0.300
History	0.073
Social Studies	0.185
Foreign Language	0.044
Arts	0.070
Special Education	0.068
Practical	0.093
Technical	0.038
Female	0.646
Race (White omitted)	
Black	0.339
Hispanic	0.013
Asian	0.005
Native American	0.022
Multiracial	0.001
Education Level (Bachelors omitted)	
Masters	0.278
Other Graduate Degree	0.016
Experience (10-19 Years omitted)	
1 Year	0.067
2 Years	0.054
3 Years	0.049
4-9 Years	0.216
20-29 Years	0.225
30-39 Years	0.070
40+ Years	0.002
Paid Bonus	0.085

Note: Sample consists of teacher/school/year observations for personnel working at least part time in a treatment or control group school. N=29,584.

Table 3: Basic Estimate of the Program's Impact

Teacher Receives a Bonus Payment	0.902* (0.053)
Teacher is Certified in Math, Science or Special Education	1.007 (0.031)
Teacher is Employed by a Currently Eligible School	1.173** (0.044)
N	29,562
Log likelihood	-56,644.23

Note: Table entries are hazard ratios, with standard errors in parentheses. The hazard refers to the probability of exiting a school after period t , conditional on remaining in that school until period t . Unit of observation is the teacher/school/year. ** denotes a hazard ratio significantly different from 1 at the 5% level; * the 10% level.

Table 4: Refined Estimates of the Program's Impact

	(1)	(2)	(3)	(4)
Teacher Receives a Bonus Payment	0.872** (0.051)	0.870** (0.051)	0.868** (0.051)	0.878** (0.051)
Teacher is Certified in Math, Science or Special Education	0.864** (0.035)	0.865** (0.035)	0.857** (0.035)	0.871** (0.036)
Teacher is Employed by a Currently Eligible School	1.167** (0.044)	1.168** (0.044)	1.155** (0.044)	1.147** (0.043)
Female	---	0.963 (0.026)	0.982 (0.027)	0.977 (0.027)
Race (White omitted)				
Black	---	1.022 (0.028)	1.021 (0.028)	1.010 (0.027)
Hispanic	---	1.230* (0.131)	1.202* (0.128)	1.194* (0.127)
Asian	---	1.169 (0.196)	1.052 (0.177)	1.064 (0.179)
Native American	---	0.721** (0.071)	0.739** (0.073)	0.701** (0.070)
Multiracial	---	1.905** (0.478)	1.786** (0.448)	1.762** (0.442)
Education Level (Bachelors omitted)				
Masters	---	---	1.132** (0.032)	1.140** (0.033)
Other Graduate Degree	---	---	1.433** (0.131)	1.454** (0.133)
Experience (10-19 Years omitted)				
1 Year	---	---	1.515** (0.068)	1.510** (0.067)
2 Years	---	---	1.274** (0.067)	1.269** (0.067)
3 Years	---	---	1.173** (0.065)	1.173** (0.065)
4-9 Years	---	---	0.931** (0.033)	0.932** (0.033)
20-29 Years	---	---	0.644** (0.025)	0.646** (0.025)
30-39 Years	---	---	1.112** (0.055)	1.116** (0.055)
40+ Years	---	---	1.566* (0.371)	1.597** (0.379)
High School	---	---	---	0.884** (0.026)
Subject Area Controls	Y	Y	Y	Y
N	29,562	29,562	29,562	29,562
Log likelihood	-56,612.39	-56,599.35	-56,418.01	-56,409.25

Note: Table entries are hazard ratios, with standard errors in parentheses. The hazard refers to the probability of exiting a school after period t , conditional on remaining in that school until period t . Unit of observation is the teacher/school/year.

** denotes a hazard ratio significantly different from 1 at the 5% level; * the 10% level.

Table 5: Testing for Heterogeneity in the Program's Impact

	(1)	(2)	(3)	(4)
Paid Bonus in 2002	0.826** (0.073)			
Paid Bonus in 2003	0.914 (0.065)	---	---	---
Paid Bonus, Teach Math	---	0.819** (0.065)	---	---
Paid Bonus, Teach Science	---	0.946 (0.080)	---	---
Paid Bonus, Teach Special Education	---	1.084 (0.144)	---	---
Paid Bonus, Teach 1 Year	---	---	0.888 (0.128)	---
Paid Bonus, Teach 2 Years	---	---	0.965 (0.185)	---
Paid Bonus, Teach 3 Years	---	---	0.958 (0.184)	---
Paid Bonus, Teach 4-9 Years	---	---	0.891 (0.099)	---
Paid Bonus, Teach 10-19 Years	---	---	0.632** (0.081)	---
Paid Bonus, Teach 20-29 Years	---	---	0.763* (0.114)	---
Paid Bonus, Teach 30-39 Years	---	---	0.959 (0.157)	---
Paid Bonus, Teach 40 Years or more	---	---	0.731 (0.548)	---
Paid Bonus, Teach in a Middle School	---	---	---	0.765** (0.078)
Paid Bonus, Teach in a High School	---	---	---	0.928 (0.061)
N	29,562	29,562	29,562	29,562
Log likelihood	-56,408.81	-56,407.66	-56,402.57	-56,407.77

Notes: Table entries are hazard ratios, with standard errors in parentheses. The hazard refers to the probability of exiting a school after period t , conditional on remaining in that school until period t . Unit of observation is the teacher/school/year.

** denotes a hazard ratio significantly different from 1 at the 5% level; * the 10% level.

All specifications include controls for being certified in Math, Science or Special Education, teaching at an eligible school, Subject taught, gender, race, experience, education and high school (with the same omissions for categorical variables as listed in Table 4).

