# Early Retirement Incentives and Student Achievement*

Maria D. Fitzpatrick
(Cornell University)

Michael F. Lovenheim
(Cornell University and NBER)

October 2012

**Abstract**

Given the prevalence of districts offering older teachers incentives to retire early in order to fill budget gaps and the rising average age of the teachers, the composition of teachers will change dramatically in coming years. However, there currently is no information on how these changes in the teacher workforce will affect student achievement. We use exogenous variation in teacher retirement behavior driven by an early retirement incentive program in Illinois in the mid-1990s to identify the effects of large-scale teacher retirements on student achievement. We find the program did not reduce test scores; likely, it increased them. The positive effects were most pronounced in low-SES and lower-performing schools. Our results suggest that districts may be able to lower costs without damaging student outcomes through early retirement programs.

KEYWORDS: Teacher Retirement, Early Retirement Incentives, Teacher Experience
JEL CLASSIFICATION: J26, I21, I28, H75

* We would like to thank Steve Rivkin, Jonah Rockoff and seminar participants at Cornell University and CESifo for helpful comments and suggestions. All errors and omissions are our own.

Maria Fitzpatrick, Department of Policy Analysis and Management, Cornell University. Mail: 103 Martha Van Rensselaer Hall, Ithaca, NY 14853. Email: maria.d.fitzpatrick@cornell.edu Phone: (607) 255-1272. Michael Lovenheim, Department of Policy Analysis and Management, Cornell University and NBER. Mail: 135 Martha Van Rensselaer Hall, Ithaca, NY 14853. Email: mfl55@cornell.edu Phone: (607) 255-0705.

## 1 Introduction

Early retirement incentives (ERI) for teachers, which offer experienced teachers financial incentives to retire before they would be eligible for full pension benefits, have become increasingly prevalent over the past decade as states and school districts seek to reduce expenditures in light of tightening budgets. While ERI programs have been in existence since the 1970s (Tarter and McCarthy 1989), their popularity has spiked in the past five years, as it has in previous recessions. For example, in 2010 alone, several large states such as New York, Michigan, and Minnesota enacted early retirement incentive legislation. California, Illinois, and Connecticut have longer standing programs, while myriad school districts throughout the country have implemented these programs in recent years. Despite the popularity of ERI programs, especially in difficult financial climates, little is known about how inducing experienced teachers to retire early affects student achievement.

The importance of examining how ERI programs affect student learning is underscored by the aging of the teacher workforce. In 2010, over one-third of teachers were over the age of fifty.[1] Given that retirement rates tend to increase dramatically around 30 years of experience, which would occur when most teachers are in their 50s, in the coming decade we can expect a large proportion of teachers to retire.[2] Indeed, in the last couple of years, many states have already seen an exodus of teachers into retirement as uncertainty around collective bargaining and cuts in retirement benefits have induced teachers to retire earlier than they otherwise would have.[3] Furthermore, virtually every state has provisions that allow teachers to retire early with an actuarially fair adjustment to their pension (Auriemma, Cooper and Smith 1992). This feature

[1] See Friedberg and Turner (2010) for a wide range of evidence on the aging of the teacher workforce.
[2] For more evidence on retirement hazard rates, see Furgeson, Strauss and Vogt (2006), Hanushek, Kain and Rivkin, (2004), and Kodel and Podgursky (2011).
[3] For example, in 2011, the Wisconsin legislature passed a bill that changed retirement contribution rates and weakened collective bargaining rights. During the first half of the year, between when the bill passed and when it would take effect, more than 5,000 public school employees retired, which is approximately equal to the number of retirees in 2009 and 2010 combined. (http://www.huffingtonpost.com/2011/08/31/doubled-teacher-retiremen_n_943495.html)

of teacher pension plans likely facilitates early retirement among many teachers.

The growing literature on teacher retirement focuses largely on the role of the retirement incentives embedded in pension plans (Furgeson, Strauss and Vogt, 2006; Costrell and Podgursky, 2009; Costrell and McGee, 2010; Brown and Laschever, 2012; Brown, 2012) and retirement's responsiveness to student performance and school demographics (Hanushek, Kain and Rivkin, 2004; Mahler 2012). Though these studies provide clear evidence that teacher labor supply is responsive to pension incentive structures, how these retirement incentives affect student outcomes is unknown.

*Ex ante*, it is unclear what the effects of large-scale teacher retirements, whether it be from ERI programs or from the aging of the workforce, will be. On the one hand, retiring teachers are highly experienced, and they typically are replaced with much less experienced teachers or with new teachers. Given the evidence of the strong relationship between experience and effectiveness in the classroom (Wiswall, 2011; Rivkin, Hanushek and Kain, 2005; Rockoff, 2004), teacher retirements could reduce student achievement due to a reduction in the experience of the workforce. On the other hand, retiring teachers may put forth less effort than younger teachers or may be less well-trained in modern, potentially more effective, pedagogical practices. This may be particularly true for those teachers retiring early. Also, principals and administrators may respond to large losses of experienced teachers by changing available resources, e.g. by decreasing class sizes or by changing the assignment of teachers to students. Resulting reduced teacher wage bills could lead to higher spending on non-teacher resources as well.

The main difficulty in identifying the effect of teacher retirement on student test scores is that retirement decisions may be endogenous to student performance. For example, teachers may retire rather than face a lower-performing cohort of students, or retirement may be induced by unobserved school-level policies that independently affect achievement. The ability for teachers to retire early exacerbates such endogeneity, as early retirement programs give teachers more

flexibility in deciding when they leave teaching.

The main contribution of this paper is to use a natural experiment in retirement incentives brought about by a two-year early retirement incentive program in Illinois in 1993-1994 and 1994-1995 to identify the effect of teacher retirement on student academic achievement. Our analysis is the first to link changes in the teacher workforce due to ERI programs to student test scores, and thus our main contribution to the literature is to provide evidence on how ERI policies impact student achievement. [4] Because we are able to link an exogenous, large number of teacher retirement decisions to student test scores, our results also yield some much needed information about how the looming retirement of the baby boom generation of teachers may affect students.

Specifically, we exploit a temporary retirement incentive that was offered to all public school teachers in Illinois. Illinois has a defined benefit pension system for teachers, where eligibility for benefits and benefit levels are defined by age and years of service in the state. In the 1993-1994 and 1994-1995 school years, the state offered an early retirement incentive, nicknamed "the 5+5," which allowed employees to purchase an extra five years of age and experience to be counted as creditable service for calculating their retirement benefit. Their purchase of these additional creditable years of service was conditional on their immediate retirement. This plan led to a threefold increase in the retirement of experienced teachers in the 1994 and 1995 school years.[5] As a result of this program, the Illinois public school system lost 10 percent of its teachers over a two year time-span, with experienced teachers making up the vast majority of exits.

Using school-level data from 1990-1997 on 3$^{rd}$, 6$^{th}$ and 8$^{th}$ grade state standardized math

[4] Mahler (2012) and McGee (2012) examine how the responsiveness of teachers to retirement incentives differs by teacher quality, but these studies do not address whether student achievement is affected by teacher retirement. Neither study finds evidence that higher or lower performing teachers are differentially responsive to performance incentives.

[5] Henceforth, we will index school years by the calendar year in which a school year ends in order to be consistent with State of Illinois practices. For example, we will refer to the 1993-1994 academic year as 1994 and the 1994-1995 academic year as 1995.

and English exams in a differences-in-differences framework, we exploit the fact that the increased retirement propensity was concentrated amongst those with the most experience in order to identify how test scores were affected by retirement. We examine how test scores changed differentially between the pre- and post-ERI periods among schools that had more teachers with 15 or more years of experience relative to those that had fewer in the pre-ERI period. We specify the strength of treatment in this manner because the data point to the program affecting the exit rates of teachers with 15 or more years of experience the most. With school fixed effects in the model, we control for any cross-sectional differences in test scores that are related to the pre-ERI experience profile in schools. The main identification assumption we invoke is that trends in student achievement among schools with fewer experienced teachers are an accurate counterfactual for trends among schools with more experienced teachers. We present evidence that this assumption is valid in this context, and thus our difference-in-difference models allow us to credibly identify the effect of teacher retirement due to the ERI program on student academic achievement.

Our estimates show that the ERI program had large effects on the composition of teachers in Illinois schools. For each teacher with 15 or more years of experience in the pre-ERI period, exit rates increased, average experience declined and the number of new teachers increased substantially post-ERI. However, despite the large literature showing a positive relationship between teacher experience and student academic achievement (e.g., Rivkin, Hanushek and Kain 2005; Rockoff 2004; Wiswall 2011), our estimates suggest that the teacher retirements did little to reduce student math and English test scores. In fact, the point estimates are positive and in some cases are statistically significantly different from zero at the 5 or 10 percent level.

The effect of teacher retirement on achievement may be heterogeneous across different types of schools and districts. For example, better funded schools may be more able to replace retiring teachers with experienced teachers from other schools.  In addition, due to teacher

sorting over their careers and the value of gaining experience in better educational settings, the experienced teachers at higher-resource schools may be higher quality. We explore the role of heterogeneity in responses across schools with different characteristics. In particular, we estimate our models for low- and non-low-income schools, by percent white, and by baseline test scores. We find suggestive evidence of heterogeneity: in lower SES and lower-performing schools, retirements from the ERI program led to larger increases in test scores, particularly for reading. We show that some of this difference is driven by how schools were affected by the ERI program, with more disadvantaged schools experiencing larger total declines in the size of the teacher workforce and at most small increases in the number of inexperienced teachers. These results fit with a story where schools differ in the quality of their existing experienced teachers before the retirement incentive program is offered (Lankford, Loeb and Wyckoff, 2002; Feng et al., forthcoming) and where less advantaged schools use the retirement program to decrease the overall number of teachers in the school.

Although we find no evidence that ERI-induced teacher retirements negatively impact student test scores, the variation used to identify our estimates differs from previous estimates on teacher experience, which generally has used individual teacher entry and exit to generate variation in experience. Instead, our estimates are driven by a large change in the composition of teachers. Furthermore, the previous literature in this area typically has been unable to identify the effects of experience for rather experienced teachers (e.g. those with more than 10 years of experience) – it could be the case that teachers who are very high in the experience distribution are less productive.[6] Finally, our estimates are identified off teachers who take up the ERI program. If these teachers are of low quality relative to the teachers who replace them, then inducing these teachers to retire could have positive effects, as we find. Although previous work has not found that responsiveness to retirement incentives differs by teacher quality (Mahler,

[6] The exception is Wiswall (2011), who estimates that the relationship between teacher experience and student test scores continues to grow throughout a teacher's career, although at a slower rate at high levels of experience.

2012; McGee, 2012), any differences between the quality of teachers retiring due to the ERI program and other teachers of the same experience level would make it difficult to use our estimates to predict the effect of increased teacher retirement due to the general aging of the teacher workforce. Our results nonetheless provide valuable information about teacher retirement and student achievement, particularly given the paucity of information on teacher retirement effects and the growing popularity of ERI programs in the US.

The rest of this paper is organized as follows: Section 2 describes the Early Retirement Incentive in the context of the pension system for teachers in Illinois. Section 3 provides a description of the data. We outline our empirical strategy and detail our results in Sections 4 and 5, respectively, before concluding in Section 6.

## 2 The Illinois Teacher Retirement System and Early Retirement Incentive

The Teachers Retirement System (TRS) in Illinois is a defined benefit pension plan, where employee contributions are made annually over the course of employment and benefits are paid out annually upon retirement. During the period studied (1990-1997), the employee contribution rate was 9 percent of earnings. Also over this period, the benefits formula was a nonlinear function of the employee's accumulated service at the time of retirement. Benefits accumulate as a percent of end-of-career salary, at a rate of 1.67, 1.9, and 2.1 percent for each year in the first, second and third decades of service, respectively, and 2.3 percent for any year of service thereafter. The maximum annual benefit employees could receive was 75 percent of their end-of-career salaries.[7]

In general, retirement benefits can be claimed by members of the TRS when they terminate active service with the Illinois Public Schools (IPS) and meet the following age and service requirements: age 55 with 35 years of service, age 60 with 10 years of service, or age 62

[7] Employees reach this maximum benefit with approximately 38 years of experience, although teachers can count up to one year of sick leave as creditable service.

with 5 years of service.[8] If a teacher is at least 55 years of age and has at least 20 years of experience, she may start collecting pension benefits, but they will be discounted by 6 percent for each year below age 60 she is. An Early Retirement Option (ERO) exists, whereby members who are at least 55 years old and who have at least 20 years of service can receive their full benefit (without the actuarial discounting of 6 percent per year) if both the employee and the employer pay a one-time fee.[9]

In 1993 and 1994, employees were offered an early retirement incentive (ERI) as an alternative to the ERO. This ERI, called the "5&5," allowed employees to buy an additional 5 years of age and service credit. As long as the member was at least 50 years old and had accumulated 5 years of service credit, she and her employer could pay a one-time fee to increase her retirement benefit as long as the teacher retired immediately: if the teacher took up the ERI, she needed to retire at the end of the 1992-1993 or 1993-1994 school years.[10] The fee for employees was 4 percent of the highest annual salary for the past five years; the fee for employers was 12 percent of the employee's highest annual salary from the last five years.[11]

Lifetime budget constraints based on the benefits formulas with and without the ERO and with and without ERI are illustrated in Figure 1. We plot the lifetime consumption for a representative teacher who was age 55 and had 20 years of experience in 1993. The vertical axis measures the present discounted value of the teacher's lifetime consumption if she retires with

[8] Teachers can retire or leave teaching prior to these age/experience levels and still receive pension benefits when they are older. For example, if a 45 year old teacher with 6 years of experience leaves IPS, she can receive the full pension benefits she has accrued when she turns 62.
[9] The ERO requires a one-time member contribution of 7 percent of salary for each year under the age of 60 or each year under 35 years of service, whichever is less. It also requires a one-time employer contribution of 20 percent of salary for each year the member is under the age of 60. Schools must allow a certain number of ERO retirements each year and pay the ERO fee if requested by the teacher.
[10] A third window at the end of the 1994-1995 school year was allowed for the small number of retirements delayed by employers.
[11] Districts had no choice but to allow teachers to retire under ERI and to make the required contribution if a teacher did. Anecdotally and in the news media, there was concern that some districts could not "afford" to allow their teachers to retire under ERI because of the direct cost (fee) to the district in doing so. However, we show the ERI program had a large effect on the exit rate of experienced teachers (see Table 3), which suggests such constraints, if they did exist, were not pervasive.

the amount of experience given on the horizontal axis. Lifetime consumption is defined as income (salary if still working, retirement benefit if retired) minus fees (if a retirement incentive was purchased). For simplicity, it is assumed that the teacher has a known age of death at 87.

The teacher's possible lifetime budget constraints before the ERI are indicated by the solid gray and black lines (for the budget constraints with and without purchasing the ERO, respectively). The rules require an actuarial discount to be made to the retirement benefit if the teacher is less than age 60 with less than 35 years of experience. Once this teacher turns 60, since she has more than 10 years of experience, this discount is no longer taken, which is what causes the kink in her budget constraint.

More relevant for our setting is the budget segment with two circular markers, which represents the budget constraint under the ERI program offered in 1993 and 1994. As the figure demonstrates, the introduction of the ERI altered the budget constraint considerably by offering the teacher the ability to retire early with higher lifetime consumption. However, the teacher must retire at the end of the 1993 or 1994 school years, which is why the change is represented as a budget segment available to her only if she retires with either 20 or 21 years of experience.

The figure shows that the 5+5 program was very generous to experienced teachers. Depending on how teachers value leisure, this program could allow for significantly higher lifetime utility than they could have attained under the existing retirement system. Although we have drawn the figure for a 55 year old teacher with 20 years of experience, a similar shift occurred for all teachers who were at least 50 years old and who had at least 15 years of experience (or who were at least 55 with 5 years of experience). The generosity of the ERI system is probably the reason it generated a significant increase in teacher exit among experienced teachers. Our identification strategy exploits this generosity to experienced teachers by using the fact that schools with more experienced teachers in the pre-ERI period were more affected by the ERI program.

## 3 Data

In various forms, the state of Illinois collects information on employees and students of its public school system. We utilize several of these data sources here. The first, the Teacher Service Record (TSR), is an administrative dataset collecting information about employees of the Illinois Public Schools (IPS). The second set of data includes school-level information on test scores for certain grades and subjects, collected since the early 1990s as part of Illinois' ongoing accountability program. Finally, we make use of information on the demographics of students in schools as reported to the Illinois State Board of Education (ISBE). We focus on the period 1990-1997, because the earliest available data are from 1990, and in 1998 the Illinois legislature changed the teacher benefit formula in ways that could influence teacher retirement decisions.[12] We describe each of these data sets in turn.

### *3.1 Illinois Teacher Service Record Data*

The TSR is collected by the ISBE. Each observation in the TSR is an employee-to-school match in a given school year. The data contain information on the number of years of creditable experience in the retirement system the employee has. In it, an employee can be followed across schools and as she enters and exits IPS. In order to focus the analysis on employees directly involved in teaching students, we subset the data to include only those staff with position titles such as "instructional staff," who are engaged as regular classroom teachers or as special education teachers. Overall, there are 852,874 teacher-year observations from 148,274 unique teachers between 1989 and 1997.

Standardized testing in Illinois is focused on $3^{rd}$, $6^{th}$ and $8^{th}$ grade students, so we restrict most of our analysis to teachers in those grades. The data include information on the lowest and highest grade served by a teacher, and we assign a teacher to a particular grade if the given grade is between the teacher's highest and lowest reported grade taught. We also observe the teacher's

[12] See Fitzpatrick (2011) for a detailed description of the 1998 policy change as well as the extent to which teachers value these benefits.

position (elementary teacher, middle school teacher, high school teacher) as well as the main subject she teaches, which includes a designation for teachers of self-contained classrooms. We assign teachers to 3rd grade if they are elementary teachers and teach a grade range that includes 3rd grade, 6th grade teachers are either elementary or middle school teachers with a grade range that includes 6th grade, and 8th grade teachers are middle school teachers who report a grade range including 8th grade. When we split teachers out by subject, we assume math teachers are those designated as "math" or "self-contained," while we assume English teachers are designated as "English/reading," "self-contained" or "bilingual."

Approximately 12% of the teacher-year observations (or 14% of teachers) do not have grade information. The majority of the missing data (90.3%) are from the City of Chicago School District, with no other single school district containing more than 0.1% of the missing grade information. Almost 54% of the teacher-year observations in Chicago are missing grade information. Some of the missing data come from schools that never report grade information for their teachers: 243 schools (5.3% of all schools in Illinois) fall into this category. We exclude teachers in these schools from our sample, which eliminates 63,344 teacher-year observations, or 11,317 teachers.[13]

We impute the missing grade information for the remaining observations without grade information. First, we leverage the longitudinal structure of the data and assign teacher to the grades they taught in other years when it was reported. This method allows us to assign a grade to 41% of the remaining missing observations. For the remaining teachers missing grade data, we assign each teacher a probability of teaching each grade based on the existing proportion of teachers in their school, year and position.[14] Using the empirical distribution of teachers in the

[13] Over 99% of these schools are in Chicago, which represent 32% of the schools in that district, and these observations represent about one half of all of the missing grade observations in the data.

[14] For example, if 30 percent of the teachers in a given school and year who are elementary schools teachers teach 3rd grade, all teachers in that school whose position is "elementary school" with missing grade information would be assigned a likelihood of 0.3 of teaching 3rd grade. When we sum teachers by grade, such teachers would count as 0.3 of a 3rd grade teacher.

school by position in this way assumes that the teachers with missing information are the same as those without with respect to grade taught. In our data, the two groups of teachers have nearly identical experience distributions, which is the only other observable characteristic we have for the teachers. As shown in Section 5.5, our results and conclusions are robust to excluding all Chicago schools as well as to excluding all imputed observations.

The final teacher-level data set we use for our analysis contains 253,463 teacher-year observations from 54,550 teachers in 3rd, 6th and 8th grades. For these teachers, we measure teacher experience using the reported total years of experience both in IPS and out of Illinois.[15] Our main measure of ERI treatment intensity, which we discuss in more detail below, is the average number of teachers in each grade (and sometimes subject) at a school in the pre-treatment period with 15 or more years of experience. The data do not include teacher age, which is why our treatment measure does not use the age-specific rules embedded in the Illinois retirement system. We take the average across the whole pre-treatment period, i.e. 1990 to 1993, to reduce the noise in our measurement of which schools should be most affected by the ERI.

Table 1 provides descriptive statistics, overall and by grade, for this variable.[16] On average, schools have 9.4 teachers, 5.2 of which have 15 or more years of experience pre-ERI. While the number of teachers increases across grades, the proportion of "experienced" teachers with 15 or more years of experience is similar, at between 52 and 57 percent. We also calculate the numbers of math and English teachers separately, as these are the subjects in which students are tested. There are about 1.5 math teachers and 1.8 English teachers on average in these three grades with 15 or more years of experience pre-ERI, which represent 56.1 and 55.4 percent of the total number of subject-specific teachers. These proportions also are similar across grades, but the standard deviations make it clear that there is much cross-sectional variation in the

[15] Including experience outside of Illinois adds little variation to overall experience, as total experience and experience in Illinois have a correlation of 0.98.

[16] While all tabulations in Table 1 are for the entire sample period, the teacher counts are calculated using only pre-1994 data. The average over the years 1990-1993 are then applied to every year.

experience composition of teachers across schools prior to the ERI program.

We also use the teacher-level data to calculate exit rates, average experience in all years and the proportion of new teachers (i.e., teachers with no more than one year of experience) in each school and year. One drawback of our data is that we do not know if a teacher took up the ERI program, *per se*. However, because we have linked teacher data over time for the entire state, we can identify the effect of the ERI program on retirement using differences in exit rates of experienced teachers when the program was implemented. It is this variation that underlies our identification strategy in which we compare test score changes in schools with more experienced teachers (that therefore would have been more affected by the ERI) to those with less experienced teachers.

### *3.2. Illinois State Board of Education Data*

We combine the administrative teacher data with data on average test scores and demographics. As part of a statewide accountability system, the IPS administers exams in math and English in grades 3, 6 and 8. We observe the average score by school, year and grade on each exam, which we scale to have a mean of zero and standard deviation of 1 in each year, grade and subject. This standardization reduces any bias associated with changes in the content or difficulty of the exam from year to year. Overall, there are 2,072 schools with a 3$^{rd}$ grade, 1,765 schools with a 6$^{th}$ grade and 797 schools with an 8$^{th}$ grade.

The ISBE data contain school demographic data as well, including percent of students who are low income, percent who are white, black, Hispanic, or Asian and Native American, and the percent limited English proficient (LEP). The school districts also record attendance rates and grade-specific enrollment for each school and year. The percent low income is defined as the percent of students receiving free- or reduced- price lunches or other Federal or state assistance. Means and standard deviations of these variables are included in Table 1.

## 4 Empirical Methodology

### *4.1. Measuring Treatment Status*

As discussed in Section 3, we do not observe ERI take-up directly. Even if we did observe take-up, however, it is likely to be endogenous to trends in student achievement (Hanushek, Kain and Rivkin, 2004). Therefore, to measure treatment status, we exploit the fact that, because of the ERI rules, teachers with 15 or more years of experience were the most likely to take up the early retirement offer.[17] This prediction is borne out in the data. In Table 2, we show exit rates in the pre-treatment years (1990-1993) and in the treatment years (1994-1995) by experience level.[18] The differences in the third column are consistent with the largest shifts in exit rates occurring among those with 15 or more years of experience. For example, for those with 15-19 years of experience and with 20-24 years of experience, the likelihood of exiting doubles. Among those with 25 to 29 years of experience, the likelihood of exit increases by around 150% relative to the baseline. The exit rate for the 30-34 years of experience group also increases substantially, from 11.1% to 36.7%. Even the exit rate among those with more than 40 years of experience jumps sharply. Thus, the amount of retirements a school experienced as a result of ERI should be directly related to the number of teachers with 15 or more years of experience in the pre-ERI period.

The second set of three columns in Table 2 shows how the ERI program affected the distribution of experience among teachers exiting. Each column shows the proportion of those exiting with a given experience level. Consistent with the fact that this program increased retirement among experienced teachers, there was a marked outward shift in the experience of exiting teachers. The final three columns of Table 2 present similar distributions for entering teachers. The largest change induced by the ERI program was a 10 percentage point (or a 19.6%)

[17] The structure of the retirement incentive in the context of the existing retirement system was such that teachers with as little as 5 years of experience became eligible to retire (as long as they were at least 50 years old). Most teachers aged 50 years or older have at least 15 years of experience, however.

[18] Note that 1993 is a pre-treatment year because the ERI induced teachers to retire at the end of this school year, which would then potentially affect achievement in the 1994 school year.

increase in the proportion of new teachers (those with one year or less of experience) among those entering the Illinois Public Schools. As expected when there is a large exodus of teachers, many replacement teachers are hired, and these new teachers are much more likely to have little to no experience. Given the negative relationship between being a new teacher and student test scores (Wiswall, 2011; Rivkin, Hanushek and Kain, 2005), this influx of inexperienced teachers is predicted to have a negative effect on student test scores, *ceteris paribus*.

Figure 2 presents additional evidence supporting the use of the number of teachers with 15 or more years of experience pre-ERI to measure treatment intensity. The figure shows probability density functions (PDFs) of teacher experience in the pre-treatment period (1990-1993) and in the two main treatment years (1994-1995). We show PDFs for each period by quartile of the proportion of teachers in grades 3, 6 and 8 with experience greater than or equal to 15 years in the pre-treatment years. As the figure demonstrates, the largest declines in experience occur among the schools with the highest pre-existing experience levels. For the low experience schools, the experience distribution shifts out slightly (as teachers gain more experience over time), while for the higher experience schools, during the treatment period, there is a large increase in the number of inexperienced teachers and a decline in highly experienced teachers. Figure 2, combined with Table 2, shows that the proportion of teachers with 15 or more years of experience before the ERI program was meaningfully related to experience changes and exit rates post-treatment. This evidence suggests that the pre-existing number of teachers with this high experience level is a strong proxy for the intensity of treatment by the ERI program.

### *4.2. Estimation Strategy*

Using the number of teachers with 15 or more years of experience pre-1994 as our measure of treatment intensity, we estimate difference-in-difference models to examine how the ERI differentially affected schools with more experienced teachers. We estimate regressions of the following form:

$$Y_{igt}^{s} = \beta_0 + \beta_1 (Teachers \geq 15)_{ig} * Post_t + \beta_2 Teachers_{ig} * Post_t + \gamma X_{it} + \delta_{ig} + \varphi_{tg} + \varepsilon_{igt}^{s} \qquad (1)$$

where $Y_{igt}^{s}$ is the standardized test score in grade *g* (3rd, 6th or 8th) for subject *s* (math or English) in school *i* and year *t*. The variable $Teachers \geq 15$ is the average number of teachers in a given grade (and in some specifications in a given grade and subject) with at least 15 years of experience pre-1994. *Teachers* is the average total number of teachers in a grade and school (or in a grade-school-subject) in the pre-ERI period and *Post* is an indicator variable equal to 1 for school years 1994 through 1997, the years during and after ERI. The vector *X* contains the set of school-by-year demographic variables discussed in Section 3 and shown in Table 1 as well as a quadratic in grade-by-year enrollment, and the model includes both school-by-grade ($\delta_{ig}$) and grade-by-year ($\varphi_{tg}$) fixed effects.[19] All regressions are weighted by average grade-specific enrollment pre-1994 so that our estimates capture the effects of the ERI on the test scores of the average student. We report standard errors that are clustered at the school-grade level to account for within school and grade serial correlation of the errors and for heteroskedasticity that typically is present in aggregate data.

The coefficient of interest in equation (1) is $\beta_1$, which is the difference-in-difference estimate of the effect of the ERI program on student achievement. The model is identified by comparing changes in test scores when the ERI is introduced across schools with fewer or more potentially affected teachers, as measured prior to ERI introduction. As discussed above, the intensity of treatment is directly proportional to the average number of teachers with 15 or more years of experience at a school in the pre-treatment period. This treatment intensity measure does not vary across school-grades over time. Therefore, any fixed differences across school-grades that are correlated with it are controlled for with the school-by-grade fixed effects. Furthermore, since we fix experience levels using pre-treatment data, the experience differences are not

[19] We take the base unit of observation in this study as the school-grade. About 56% of the schools have two grades, and 12% of the schools contain 3rd, 6th and 8th grades.

endogenous responses to the treatment itself.[20]

The main identifying assumption in equation (1) is the following: absent the ERI program, school-grades with different experience levels would have had the same trends in student test scores, *ceteris paribus*. That is, conditional on school-grade fixed effects and time-varying observable characteristics, trends in achievement among school-grades with a low proportion of experienced teachers provide an accurate counterfactual for trends among school-grades with a high proportion of experienced teachers.[21] As we demonstrate in Section 5.2., there is no evidence that school-grades with more highly experienced teachers are trending differently than those with fewer experienced teachers in the pre-treatment period. Because we know of no other reason test scores would change differentially by the pre-treatment level of experience when the ERI program comes into effect, we believe $\beta_1$ provides a credible estimate of how the teacher retirements that occurred due to the ERI program impacted student achievement.

## 5 Results

### *5.1. The Effect of ERI on Teacher Exits, Experience and Student-Teacher Ratios*

Before examining how the teacher retirement incentive program affected student academic achievement, it is important to understand how it impacted schools through changes in retirements, teacher experience and school resources. In Table 3, we present estimates of equation (1), where the dependent variable is each of the following in turn: the number of experienced teachers (i.e., teachers with 15 or more years of experience) in a school-grade who exit in the previous year, average experience in the school-grade, the proportion of new teachers

[20] Note that because neither *Teachers* nor $Teachers \geq 15$ vary within school-grade over time, they are perfectly collinear with the school-grade fixed effects and thus are omitted from equation (1).

[21] Recall that we control for the number of teachers, so the model is essentially identified using, in part, cross-sectional differences in the proportion of experienced teachers. Models that use this proportion directly yield similar results, which are shown in Section 5.5. But, we prefer controlling for the number of total teachers and the number of teachers with at least 15 years of experience separately as it is more flexible than the model that uses the proportion of experienced teachers.

in the school-grade, and student-teacher ratios. These estimates describe how the characteristics of the teacher workforce changed due to ERI as a function of the number of teachers with 15 or more years of experience pre-ERI.

Table 3 shows the results from estimation of equation (1) with these dependent variables, both pooled across grades as well as by grade. Each cell in the first column comes from a separate regression, while the estimates in each row of the final three columns come from one regression that includes interactions between $Post*Teachers \geq 15$ and grade indicators as well as *Post*Teachers* and grade indicator interactions. The table presents the difference-in-difference estimates only, although all models contain the full set of controls shown in equation (1). Table 3 indicates that the ERI program had a large impact on experienced teacher exits and on experience, which is consistent with Figure 2 and Table 2. Having one more teacher with 15 or more years of experience pre-ERI increased the number of experienced teachers (i.e., teachers with at least 15 years of experience) exiting post-1993 by 0.08 overall, with estimates that vary little across grades. Relative to the baseline, this represents a 32.6% increase in the experienced teacher exit rate for each teacher over 15 years of experience pre-ERI.

The increased exit among more experienced teachers was accompanied by declines in average teacher experience, with experience declining by 0.41 years, or 2.6%, for every teacher with 15 or more years of experience pre-ERI. The experience effects are largest in 3rd grade, with a 1.1 year, or 7.1%, decline in average experience for each teacher with 15 or more years of experience. As predicted by Figure 2, the drop in average experience was driven in some part by an increase in the number of new teachers. The third panel of results shows the effect of ERI on the number of new teachers. Overall, for each teacher pre-ERI with 15 or more years of experience, the number of new teachers increases by 0.073, or 19.8% relative to baseline. The first year teacher effect is largest in levels in grade 6, at 0.088, but relative to baseline it is largest in 3rd grade, with a 48.7% increase in new teachers for each teacher with more than 15 years of

experience. All of these estimates are statistically different from zero at the 5 percent level, and they suggest that school-grades with more teachers with 15 or more years of experience pre-ERI experienced much larger changes in teacher turnover and teacher experience after the program was implemented.

Table 4 provides a more complete accounting of how the experience composition of the teacher workforce changed post-ERI as a function of the number of teachers with 15 or more years of experience. The table shows estimates from models akin to equation (1) but that use the number of teachers in each experience group as a dependent variable. As with Table 3, we estimate average effects across grades as well as separately by grade by including grade interactions. The estimates point to a downward shift in experience that is driven by an increase in the number of teachers with less than 15 years of experience and a decrease in the number of teachers with more than 15 years of experience. The largest increases are in teachers with under 10 years of experience, and the largest declines are in teachers with 15-29 years of experience. Furthermore, these changes are relatively similar across grades. Summing all of the estimates across grades in a column gives the impact of having one more experienced teacher pre-ERI on the total number of teachers post-ERI. In the pooled model, this estimate is -0.289, which suggests that the teacher workforce shrinks in schools with more experienced teachers due to this program. This average reduction is driven entirely by the 6$^{th}$ and 8$^{th}$ grades, which shrink by -0.265 and -0.364 teachers, respectively, for each teacher with more than 14 years of experience pre-ERI. The 3$^{rd}$ grade teacher workforce is basically unchanged, with a slight increase of 0.02 teachers for each pre-ERI experienced teacher. Thus, at least for the two higher grades, schools appear to be using this program to reduce the total number of teachers in the school, perhaps because the inability to fire teachers under union contracts makes it difficult to shrink the workforce in other ways.

While retirements under the ERI program occurred rather quickly and unexpectedly, the

possibility remains that administrators altered other education inputs to compensate for the loss of experienced teachers. Such offsetting behavior would be of interest in its own right, and we stress that our test score results below are net of any of these endogenous adjustments. Unfortunately, detailed data on curriculum, resources and expenditures are not reported in ways that would prove useful for measuring offsetting behavior by principals and teachers. One potentially important input we can observe, however, is pupil-teacher ratios. In the last row of Table 3, we estimate equation (1) using pupil-teacher ratios as the dependent variable. For all grades combined, we find a statistically significant effect of -0.032, which is less than a 0.2 percent change per teacher with 15 or more years of experience. The pupil-teacher ratio effects are largest in absolute value in 3rd grade, at -0.076, and decline as the grades increase such that the 8th grade estimate is -0.014 and is not statistically significant at even the 10% level. However, even the 3rd grade estimate is very small relative to baseline, much smaller than the decreases in class sizes that have shown positive effects in the literature.[22]

Together, Tables 3 and 4 point to large changes in the composition of teacher experience post-ERI in schools that had more experienced teachers pre-ERI. We now turn to an examination of how test scores were affected by these changes.

### *5.2. The Effect of ERI on Student Achievement*

Although equation (1) constitutes our preferred empirical model, we first present "event study" estimates of the effect of the ERI on student test scores in order to examine whether there are pre-treatment trends in scores as a function of the pre-treatment experience distribution and whether there are time-varying treatment effects that our main specification might miss. Figure 3 presents these estimates for the pooled sample using all grade-specific teachers, and Figure 4 presents estimates that use the number of grade and subject-specific teachers. The results come

[22] There is a large literature on how class sizes affect student achievement, but there is little consensus on whether they do and on how large any effects are. Studies examining the Tennessee STAR class size experiment (e.g., Krueger, 1999) tend to find large effects, while Hoxby (2000) finds no effect of class size on achievement using class size discontinuities driven by maximum class size rules and population variation. See Hanushek (2003) for a critical review of this literature.

from models similar to equation (1), except the $Post*Teachers \geq 15$ and *Post*Teachers* have been replaced with year indicators interacted with $Teachers \geq 15$ and *Teachers.* The figures show the $Teachers \geq 15 *I(year)$ coefficients along with 95% confidence intervals that are calculated from standard errors that are clustered at the school-grade level. The 1993 coefficient is set to zero by design, so all estimates are relative to this year.

In Figure 3, there is no evidence of any differential trend in test scores as a function of the number of teachers with 15 or more years of experience prior to the implementation of ERI. As discussed in Section 4, our difference-in-difference estimates are identified under the assumption that test scores in schools with different pre-treatment experience levels would have trended the same, and the estimates in Figure 3 support this assumption. Furthermore, after 1993, the figure indicates that test scores actually rose. While there is a small decline in math in 1994, on the order of -0.005 standard deviations, it is not statistically significantly different from zero.[23] Test scores then increase for the next several years with the number of pre-treatment experienced teachers. For reading, all of the post-ERI estimates are positive, and they are statistically significantly different from zero at the 5% level for all years except 1995.

Figure 4 presents the same results but using subject-specific teachers. Again, there is no evidence of pre-treatment trends, and the post-ERI estimates are mostly positive. However, these estimates are less precise than the estimates in Figure 3, so the coefficients typically are not statistically different from zero. However, the findings here are very similar to those in Figure 3 and not only support our identification strategy but also suggest that the changes to schools shown in Tables 3 and 4 brought about by ERI-induced retirements did not reduce student test scores.

---

[23] Appendix Figure A-1 contains event study estimates for teacher exits, average experience and the number of new teachers. For each outcome, there is a sizable shift in 1994 and 1995 that is not predicted by pre-treatment trends. Furthermore, the effects on teacher exit and experience are larger in 1995 than in 1994. If the small decline in math test scores in 1994 were being driven by changes in teacher turnover and experience in that year, we would expect to see an even larger negative effect in 1995, which is inconsistent with the estimates in Figure 3.

Figures 3 and 4 suggest that our difference-in-difference model given by equation (1) is not biased by the existence of differential pre-treatment trends as a function of pre-ERI experience, but due to the demanding nature of these models the program effects shown in the figures are rather imprecise. We now turn to estimates of equation (1) in order to increase the precision of the results relative to the event study model. Baseline estimates of equation (1) are shown in Table 5. Each column represents a separate regression, with the first two columns showing estimates using all grade-specific teachers and the second two showing results using subject-grade-specific teachers.

The results are consistent with Figures 3 and 4, indicating that test scores increase post-ERI among school-grades that have more teachers with 15 or more years of experience pre-ERI. For all but math in the first column, these estimates are statistically significantly different from zero at the 5% or 10% level and are about 0.01 standard deviations in magnitude. For math in the first column, the estimate is smaller, at 0.003 standard deviations and is precise enough to rule out a negative effect smaller than -0.0035 standard deviations at the 5% level. In the third column, we can rule out an effect smaller than -0.0018 at the 5% level. Thus, despite the fact that having one more experienced teacher pre-ERI leads to significant declines in the teacher experience profile and significant increases in exit at the school-grade level, test scores rise, or at the very least do not decline.

Although Table 3 suggests that the ERI program had similar effects on teacher composition across grades, the impacts on some measures, such as average experience, are larger in grade 3. In Table 6, we estimate equation (1) including interactions between *Post*Number* $\geq$ *15* and grade indicators as well as *Post*Teachers* and grade indicators. While the estimates become much less precise, they are, on the whole, qualitatively and quantitatively similar to the pooled estimates. There are a couple exceptions, however. First, there is now a negative effect for reading in both specifications for 3rd grade. But, the point estimates still are small in absolute

value and are not statistically significant at conventional levels. Second, using subject-specific teachers, there are large positive effects in $8^{th}$ grade in both subjects that are significant at the 10% level. While these estimates are somewhat noisier than the pooled estimates, they still provide little evidence of a decline in test scores due to the ERI-induced teacher retirements.

Given that the average retiring teacher had 29 years of experience post-ERI and the average incoming teacher had less than 3 years of experience, our finding that these retirements had little effect on student achievement is puzzling in light of much of the literature on teacher experience. For example, Rivkin, Hanushek and Kain (2005) show that new teachers perform about 0.07-0.13 standard deviations worse in math and 0.03-0.06 standard deviations worse in English. Rockoff (2004) presents estimates that generally are of the same magnitude, with some evidence of a much larger experience impact for reading comprehension. In a review of the literature, Wiswall (2011) reports a host of previous studies that find new teachers perform between 0.10 and 0.18 standard deviations worse than experienced teachers. He then estimates returns to experience that are an order of magnitude larger using non-parametric methods. Since we find little support for the notion that schools are shifting resources in order to counteract the potential negative impact of teacher retirements (see Section 5.5), we suspect that the teachers who took up the ERI were less productive teachers than the ones that replaced them or than the ones remaining in the school.[24] This finding is in contradiction as well to previous work showing high and low quality teachers are similarly responsive to pension incentives (McGee, 2012; Mahler, 2012).

### *5.3. Heterogeneous Treatment Effects*

Thus far, our estimates suggest that the teacher retirements induced by the ERI program did not reduce student achievement, and they may have increased test scores, on average.

[24] It is also possible that the previous literature contains upward biases driven by the endogeneity of teacher turnover, which is the main source of variation used. However, the estimates in Wiswall (2011), which use within teacher variation in experience to identify its effects on teacher quality, do not suffer from this upward bias and would have predicted ERI to have large negative effects.

However, there could be heterogeneous treatment effects driven, at least in part, by the fact that wealthier or higher-achieving schools may find it easier to replace retiring teachers with experienced teachers from other schools. If such heterogeneity exists, it suggests any negative effects of retirement on lower-SES and lower-achieving districts would be larger. On the other hand, since teachers tend to move to higher-achieving and higher-SES schools as they gain experience (Lankford, Loeb and Wyckoff, 2002), low SES/achievement schools may not have had much turnover from early retirement if their teacher workforce were too young to take advantage of the program. Furthermore, retirements may be altering the size of the teacher workforce differentially in different types of schools, which could lead the program to have different impacts on test scores.

In Table 7, we estimate equation (1) separately by percent low income, percent white and pre-treatment average test scores. Across all three margins, the "lowest" group in the table refers to the bottom quarter of the pre-ERI distribution, while the "other" category is the top 75 percent.[25] For all but one estimate in Table 7, the coefficients are positive, and the negative estimate for math for low income schools is only -0.001. Although the estimates are somewhat imprecise, especially for the low resource and low performing schools, a general pattern emerges from the table. For reading, the low income, high minority and lowest baseline schools experience larger increases in test scores post-ERI for each teacher with 15 or more years of experience pre-ERI than their wealthier, low-minority, higher-performing counterparts. In math, such a difference only is evident for the lowest versus highest baseline score schools.

Table 7 points to general increases in student achievement from teacher retirement that are the largest for disadvantaged schools. The differences across school types we find could be driven by the fact that teachers tend to move to higher-performing and wealthier districts over

[25] More specifically, we calculate the average of each of these variables by school in the pre-treatment period and assign schools to low and high groups based on this distribution. Thus, assignment is a fixed characteristic of a school and does not change over time within school.

the course of their careers and that this type of mobility is easier for better teachers. Alternatively, Feng et al. (forthcoming) shows that the return to experience in Florida and North Carolina is lower in high-poverty relative to low-poverty schools, perhaps because of differences in the quality of the environment in which teachers gain experience. Either story is consistent with the implication of our finding that teachers who remain in poorer schools at the end of their careers are lower-performing than the teachers who replace them.

Another reason for the heterogeneity shown in Table 7 is that the ERI may have had different effects on retirement behavior and experience in different school types. For example, if low-income schools were only able to hire novice teachers to replace retirees but high-income schools hired mid-career teachers, it would help to reconcile our results with the previous literature on teacher experience and teacher quality. Alternatively, rather than replacing experienced teachers with novices, lower-resource and under-performing schools may use the ERI as a way to shrink their teacher workforce. To see if this is the case, in Table 8, we present estimates by school type of the effect of this program on the number of experienced teachers exiting, average experience, the number of new teachers and pupil-teacher ratios.

Focusing first on exits, lower income, high minority and lower-performing schools experience less turnover due to ERI, although the point estimates still are sizable relative to the baseline exit rates. Consistent with this difference in exits, less-advantaged schools experience smaller declines in experience, even though having more experienced teachers pre-ERI leads to a statistically significant decline in average experience in all school types. The starkest difference across school types is in the effect of the program on the number of new teachers. In the low-SES and low-performing schools, the number of new teachers declines post-ERI with the number of teachers with 15 or more years of experience pre-ERI. The ERI program led to large increases in the number of new teachers, particularly relative to baseline, in other schools.

That the number of new teachers declines as a function of pre-ERI experience in lower-

SES and lower-performing schools is somewhat surprising. In order to shed some light onto how the ERI program impacts the distribution of teacher experience and the number of teachers in these schools, Appendix Table A-1 contains similar estimates to Table 4 but for the lowest quartile of schools in the distributions of income, percent minority or baseline scores. In all four columns, having more pre-ERI experienced teachers leads to a significant and large decline in the number of teachers with 15-30 years of experience. However, unlike the estimates in Table 4, there are only small increases in the number of mid-career teachers (those with 4 to 14 years of experience), and the number of new teachers with less than 4 years of experience actually declines. The total number of teachers in these schools declines significantly, by between -0.38 and -0.50 for each teacher with 15 or more years of experience pre-ERI.[26] Thus, average experience declines in these schools because experienced teachers retire, but unlike for the more advantaged schools, they are not being replaced at the same rate by less experienced teachers. Instead, the teacher workforce in these schools is shrinking. Consistent with this hypothesis, student-teacher ratios rise as a function of the number of pre-ERI experienced teachers in the lower performing and lower resource schools, as shown in Table 7. Larger test score effects in the more disadvantaged schools are likely driven by some combination of less productive teachers retiring and a decline in the number of new teachers. Our data and empirical design do not let us distinguish between these two theories, however.

### *5.5. Potential Mechanisms*

In addition to the retirement of older teachers who are replaced by less experienced teachers, there are several alternative mechanisms that may contribute to the positive effects of the ERI-induced retirements on test scores. These mechanisms are not sources of bias, but rather, are some of the ways through which early retirement incentive programs, or teacher retirement more generally, might lead to increases in student achievement. Though our baseline estimates

[26] The effect of ERI on the total number of teachers can be found by summing the coefficients in Table A-1 in each row.

identify the effect of the ERI program net of any of these mechanisms, which is a policy effect of interest, it is informative to explore some of the ways in which teacher retirement might impact student achievement other than through altering the experience profile of the teacher workforce.

First, as noted above, principals or administrators could shift resources towards the most affected grades to reduce the impact of teacher turnover. However, as shown in Table 3, student-teacher ratios decline only slightly as a function of the number of pre-ERI teachers with 15 or more years of experience. Alternatively, school expenditures may increase in the schools that lost the most teachers if the school district finance allocation method does not fully adjust for the change in the teacher wage bill. Unfortunately, no information exists on school-level budgets during this time period in Illinois, but previous work has not found consistent evidence that higher spending increases student academic performance.[27]

We examine the role of resources by estimating the effect on test scores of teachers retiring in other grades in the school. If retirement leads to an increase in school resources, this effect is unlikely to be restricted to own-grade retirements only. In Table 9, we show estimates of equation (1) in which we include the number of teachers and the number of teachers with 15 or more years of experience in higher grades pre-ERI in each school. We use only higher-grade teachers because lower-grade teachers can influence current student test scores. This method necessitates that we exclude 8th grade, as there are very few schools with 8th grades that have higher-grade teachers in them. The first two columns of Table 9 show that excluding 8th grade attenuates the estimate on the number of teachers with 15 or more years of experience, as suggested by Table 6, but that we obtain the same qualitative result. When we include other grade teachers, the estimate on the number of other-grade teachers with 15 or more years of experience pre-ERI is negative, but is not statistically significantly different from zero. These

[27] See Hanushek (2003) for a critical review of this literature. Hoxby (2001) also shows evidence that spending increases brought about by school finance reforms had no effect on dropout rates, but Papke (2005) finds that school finance reform in Michigan increased 4th grade test pass rates.

results provide evidence against the hypothesis that teacher retirements led to higher school resources, which then increased test scores.

A second potential mechanism driving our results is a change in teacher assignments post-ERI. With the influx of new teachers, these teachers could be assigned disproportionately to non-tested grades. In results available upon request, we have found no evidence that new teachers were more likely to sort into grades 3, 6 or 8 post-ERI relative to the period before the program was introduced. However, existing teachers could have moved into tested grades from non-tested grades. Such an increase in experienced teachers could raise student test performance. In Figure 5, we show experience distributions of teachers who taught in Illinois both before and after ERI was introduced by whether they switched across tested and non-tested grades when the ERI program was implemented. For each teacher, we use the maximum level of pre-ERI experience to specify the experience level. As Panel A of the figure demonstrates, there was a sizable shift among mid-career teachers into tested grades from untested grades. The increase was largest among those with 5-7 years of experience pre-ERI. However, the second two panels of the figure suggest these changes in teacher assignment were not strongly related to teacher retirements. These panels show the experience distributions by switching status among teachers in schools in the bottom and top quartiles of the percent of teachers with 15 or more years of experience pre-ERI. Though there was an increase in the top quartile of schools of early-mid career teachers switching into tested grades, the increase was larger in the lowest quartile of pre-ERI experience, where the treatment effects were smallest. Thus, there is some shifting of teacher assignments among schools in response to the outflow of experienced teachers that led to more mid-career teachers in tested grades. But, these changes in the most highly treated schools are small, and it is unlikely that they are of sufficient magnitude to drive the positive effects on test scores we find.

### *5.5. Robustness Checks*

In this section, we explore the robustness of our estimates to several of the modeling and data assumptions made throughout the analysis. Table 10 shows a series of robustness checks that demonstrate how our estimates change with these assumptions. Each cell is a separate regression, and we show estimates using all grade-level teachers as well as using subject-specific teachers in the grade. In the first row of Table 10, we measure the intensity of ERI treatment with the number of teachers with 20 or more years of experience. We do this check because of the small number of retirements among teachers with 15-19 years of experience (see Table 2). These estimates are very similar to those in Table 5. The only difference is that the subject-specific estimates are slightly smaller and are not statistically different from zero at conventional levels.

Next, instead of specifying treatment intensity using the number of teachers with 15 or more years of experience (controlling for the total number of teachers), we use the percentage of teachers with 15 or more years of experience (without controlling for the total number of teachers). This functional form change in the treatment variable makes it difficult to compare the estimates with the baseline ones, but an increase in one experienced teacher is a 0.11 percentage point increase in the proportion of experienced teachers overall, and it is a 0.38 and 0.31 percentage point increase for math and English, respectively. Multiplying the estimates in the second row of Table 10 by these percentages produces estimates that are very similar to baseline for the models using all teachers and that are somewhat smaller though still qualitatively similar using the subject-specific teachers. The main finding of this analysis that teacher retirements due to ERI did not decrease (and may have increased) test scores is unchanged when measuring treatment intensity using the proportion, rather than the number, of experienced teachers.

In the final two rows, we examine the robustness of our estimates to our treatment of missing grade information for teachers in the data. Because the vast majority of the missing data are from Chicago, we first drop all Chicago schools from the analysis. Then, we drop all teachers with missing grade information and redo our estimates. Both sets of results are virtually identical

to the baseline results, which suggests our imputation of the missing grade information is not driving our results and conclusions.

## 6 Conclusion

This paper presents the first evidence in the literature on the effects of teacher retirement in response to early retirement incentives on student achievement. We use an exogenous increase in the incentive to retire among Illinois public school teachers in 1994 and 1995 that induced large numbers of teachers to retire. Using the fact that schools that had more experienced teachers prior to the implementation of the program were more affected by it, we use a difference-difference framework to determine the effects of the ERI on student achievement.

Although we show that the ERI program led to a large amount of retirement by experienced teachers, which consequently lowered teacher experience levels, we find the program did not reduce test scores and instead led to increased student achievement in most cases. Our estimates are sufficiently precise to rule out even small negative effects of the program on math and English scores. We also show suggestive evidence that the ERI program had larger positive effects in more disadvantaged schools.

These results raise a puzzle of why test scores would increase when large numbers of experienced teachers retire and why this effect would be larger for lower-SES and lower-performing schools. One potential explanation is that the returns to experience are negative high in the experience distribution. If this is true, replacing senior teachers with novice ones would increase test scores. This effect could be particularly large for less advantaged schools, as more effective teachers tend to move to higher-performing and higher-SES schools early in their career (Boyd et al, forthcoming; Lankford, Loeb and Wyckoff, 2002). Such a pattern of selection over time could lead to the least effective experienced teachers being located in the lowest resource schools.

Alternatively, it could be the case that there is adverse selection in who responds to the ERI opportunity. If the lowest-quality teachers are those whose employment is most elastic with respect to retirement incentives, test scores would increase as these teachers are replaced with newer ones. Again, perhaps career mobility patterns have left lower-income and lower-performing schools with teachers of lower quality, rendering the effects in these schools the largest. In such a case, our estimates yield information on the effect of ERI programs on student achievement, but it would be more difficult to use them to predict the effects of the large impending volume of teacher retirements due to the aging of the teacher workforce.

A final explanation for which we find support in the data has to do with shrinking of the workforces in lower-SES schools. Because these schools did not hire inexperienced teachers when the experienced teachers left, their workforces shrank. Test scores in these schools may have grown more than those in higher-SES schools because of the lack of influx of very inexperienced teachers. Given available data, it is difficult to examine empirically the relative importance of these stories, but the implication of our results is clear: offering expiring incentives for late career teachers to retire does not harm student achievement on average.

From a broader policy perspective, our estimates suggest ERI programs could be beneficial for school districts. Newer teachers are considerably less expensive for districts than older teachers.[28] Without a deleterious impact of retirement on student achievement, replacing older teachers with younger ones achieves savings at no cost to student test scores, which may be a desirable policy for districts seeking to fill budget gaps. To the extent that the Illinois experience is generalizable to the rest of the United States, our results thus point to the potential to reduce district costs without harming student achievement by offering teachers incentives to retire early. This type of policy may be particularly effective in the lowest-performing and lowest-SES schools.

[28] In the U.S., the mean salary of a teacher with less than 5 years of experience is $32,000, while the mean salary for one with more than 30 years of experience is $49,000 (Noel, 2010).

## REFERENCES


Auriemma, Frank V., Bruce S. Cooper and Stuart C. Smith. 1992. "Graying Teachers: A Report on State Pension Systems and School District Early Retirement Incentives." ERIC Clearinghouse on Educational Management: https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/3271/graying.pdf?sequence=1.

Boyd, Donald, Pam Grossman, Hamilton Lankford, Susanna Loeb and James Wyckoff. Forthcoming. "Who Leaves? Teacher Attrition and Student Achievement." *Economics of Education Review*.

Brown, Kristine M. and Ron A. Laschever. 2012. "When They're Sixty-Four: Peer Effects and the Timing of Retirement." *American Economic Journal: Applied Economics* 4(3): 90-115.

Brown, Kristine. 2012. "The Link between Pensions and Retirement Behavior: Lessons from California Teachers." Mimeo.

Costrell, Robert M. and Josh B. McGee. 2010. "Teacher Pension Incentives, Retirement Behavior, and Potential for Reform in Arkansas." *Education Finance and Policy* 4(4): 492-518.

Costrell, Robert M. and Michael Podgursky. 2009. "Peaks, Cliffs, and Valleys: The Peculiar Incentives in Teacher Retirement Systems and Their Consequences for School Staffing." *Education Finance and Policy* 4(2): 175-211.

Feng, Li, David Figlio, Jane Hannaway, Tim Sass and Zeyu Xu. Forthcoming. "Comparison of the Value Added of Teachers in High-Poverty Schools and Teachers in Lower-Poverty Schools." *Journal of Urban Economics*.

Fitzpatrick, Maria D. 2011. "How Much Do Teachers Value Their Pension Benefits?" Mimeo.

Friedberg, Leora and Sarah Turner. 2010. "Labor Market Effects of Pensions and Implications for Teachers." *Education Finance and Policy* 5(4): 463-491.

Furgeson, Joshua, Robert P. Strauss and William B. Vogt. 2006. "The Effects of Defined Benefit Pension Incentives and Working Conditions on Teacher Retirement Decisions." *Education Finance and Policy* 1(3): 316-348.

Hanushek, Eric A. 2003. "The Failure of Input-based Schooling Policies." *The Economic Journal* 113(485): F64-F98.

Hanushek, Eric A., John F. Kain and Steven G. Rivkin. 2004. "Why Public Schools Lose Teachers." *Journal of Human Resources* 39(2): 326-354.

Hoxby, Caroline M. 2000. “The Effects of Class Size on Student Achievement: New Evidence from Population Variation.” *Quarterly Journal of Economics* 115(4): 1239-1285.

Hoxby, Caroline M. 2001. “All School Finance Equalizations are Not Created Equal.” *Quarterly Journal of Economics* 116(4): 1189-1231.

Koedel, Cory and Michael Podgursky. 2011. “Teacher Pension Systems, the Composition of the Teaching Workforce, and Teacher Quality.” University of Missouri, Department of Economics Working Paper 11-09.

Krueger, Alan B. 1999. “Experimental Estimates of Education Production Functions.” *Quarterly Journal of Economics* 114(2): 497-532.

Lankford, Hamilton, Susanna Loeb and James Wyckoff. 2002. “Teacher Sorting and the Plight of Urban Schools: A Descriptive Analysis.” *Education Evaluation and Policy Analysis* 24(1): 37-62.

Mahler, Patten. 2012. “Retaining a High Quality Teaching Workforce: The Effects of Pension Design.” Working Paper, University of Virginia.

McGee, J.B. 2012. Who Leaves and Who Stays: An Analysis of Teachers’ Behavioral Response to Retirement Incentives. Working Paper.

Noel, A.M. 2010. Documentation for the NCES Common Core of Data Teacher Compensation Survey (TCS) Agency-Level Public-Use Data File, School Year 2006–07 (NCES 2010-356). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved [April 12, 2012] from http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2010356.

Papke, Leslie E. 2005. “The Effects of Spending on Test Pass Rates: Evidence from Michigan.” *Journal of Public Economics* 89(5-6): 821-839.

Rivkin, Steven G., Eric A. Hanushek and John F. Kain. 2005. “Teachers, Schools, and Academic Achievement.” *Econometrica* 73(2): 417-458.

Rockoff, Jonah E. 2005. “The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data.” *American Economic Review Papers and Proceedings* 94(2): 247-252.

Tarter, Scott E. and Martha M. McCarthy. 1989. “Early Retirement Incentive Programs for Teachers.” *Journal of Education Finance* 15(2): 119-133.

Wiswall, Matthew. 2011. “The Dynamics of Teacher Quality.” SSRN Working Paper. Available at SSRN: http://dx.doi.org/10.2139/ssrn.1911309.

**Table 1. Descriptive Statistics**

| Variable | Full Sample Mean | Full Sample SD | 3rd Grade Mean | 3rd Grade SD | 6th Grade Mean | 6th Grade SD | 8th Grade Mean | 8th Grade SD |
|---|---|---|---|---|---|---|---|---|
| Number With 15 or More Years of Experience | 5.19 | 5.24 | 2.27 | 1.48 | 5.07 | 4.57 | 9.76 | 6.40 |
| Total Number of Teachers in Grade | 9.38 | 8.35 | 4.33 | 2.33 | 9.36 | 7.15 | 17.01 | 9.75 |
| Number of Math Teachers in Grade with 15 or More Yrs. Exp. | 1.48 | 1.16 | 1.59 | 1.02 | 1.47 | 1.31 | 1.33 | 1.13 |
| Total Number of Math Teachers in Grade | 2.64 | 1.61 | 2.95 | 1.42 | 2.61 | 1.84 | 2.24 | 1.43 |
| Number of English Teachers in Grade with 15 or More Yrs. Exp. | 1.80 | 1.36 | 1.60 | 1.02 | 1.78 | 1.42 | 2.14 | 1.64 |
| Total Number of English Teachers in Grade | 3.25 | 1.95 | 2.98 | 1.45 | 3.20 | 2.06 | 3.71 | 2.33 |
| Percent Black | 12.80 | 24.14 | 12.60 | 23.36 | 14.91 | 27.04 | 10.20 | 20.55 |
| Percent Hispanic | 6.26 | 13.28 | 6.53 | 13.89 | 6.98 | 14.69 | 4.87 | 9.75 |
| Percent Asian | 2.66 | 4.89 | 2.71 | 4.76 | 2.67 | 5.05 | 2.56 | 4.84 |
| Percent Limited English Proficient | 2.97 | 7.43 | 3.56 | 8.53 | 3.24 | 7.88 | 1.72 | 4.18 |
| Percent Low Income | 24.26 | 23.24 | 25.04 | 23.85 | 26.96 | 25.48 | 19.40 | 17.72 |
| Attendance Rate | 95.14 | 1.66 | 95.46 | 1.42 | 95.05 | 1.76 | 94.80 | 1.75 |
| Enrollment | 107.79 | 75.53 | 74.85 | 41.67 | 104.67 | 72.82 | 161.62 | 87.64 |
| Number of Observations | 32,291 | | 14,957 | | 11,743 | | 5,591 | |
| Number of School-Grades | 4,634 | | 2,072 | | 1,765 | | 797 | |

Source: Illinois Teacher Service Record data from 1990 through 1997 combined with school-level Illinois State Board of Education data on test scores. Notes: Means and standard deviations are based on school-by-year level data for 3rd, 6th and 8th grade teachers. Teachers who teach multiple grades are included in the tabulations for each grade in which they teach. Teachers who teach in self-contained classrooms are assumed to teach both math and English.

**Table 2. Entry and Exit Likelihoods by Experience Level, Before and During the Early Retirement Incentive Program**

| Years of Experience | P(Leave) | | | Experience Distribution of Exit | | | Experience Distribution of Entry | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1990-1993 | 1994-1995 | Difference | 1990-1993 | 1994-1995 | Difference | 1990-1993 | 1994-1995 | Difference |
| ≤1 | 0.091 | 0.066 | -0.025 | 0.080 | 0.035 | -0.045 | 0.511 | 0.611 | 0.100 |
| 2 | 0.066 | 0.060 | -0.006 | 0.055 | 0.030 | -0.025 | 0.096 | 0.090 | -0.006 |
| 3 | 0.060 | 0.057 | -0.003 | 0.045 | 0.030 | -0.015 | 0.065 | 0.062 | -0.003 |
| 4 | 0.053 | 0.045 | -0.008 | 0.038 | 0.022 | -0.016 | 0.044 | 0.036 | -0.008 |
| 5 | 0.056 | 0.060 | 0.004 | 0.038 | 0.026 | -0.012 | 0.033 | 0.030 | -0.003 |
| 6-9 | 0.042 | 0.045 | 0.003 | 0.105 | 0.067 | -0.038 | 0.108 | 0.080 | -0.028 |
| 10-14 | 0.031 | 0.040 | 0.009 | 0.104 | 0.065 | -0.039 | 0.062 | 0.041 | -0.021 |
| 15-19 | 0.024 | 0.044 | 0.020 | 0.104 | 0.086 | -0.018 | 0.039 | 0.023 | -0.016 |
| 20-24 | 0.032 | 0.062 | 0.030 | 0.128 | 0.138 | 0.010 | 0.023 | 0.016 | -0.007 |
| 25-29 | 0.051 | 0.129 | 0.078 | 0.112 | 0.201 | 0.089 | 0.013 | 0.007 | -0.006 |
| 30-34 | 0.111 | 0.367 | 0.256 | 0.117 | 0.225 | 0.108 | 0.005 | 0.003 | -0.002 |
| 35-39 | 0.253 | 0.407 | 0.154 | 0.062 | 0.067 | 0.005 | 0.001 | 0.001 | 0.000 |
| ≥40 | 0.281 | 0.362 | 0.081 | 0.012 | 0.010 | -0.002 | 0.000 | 0.001 | 0.001 |

Source: Illinois Teacher Service Record data from 1990 through 1995 combined with school-level Illinois State Board of Education data on test scores.
Notes: All years refer to the calendar year in which a school year ends and are based on all teachers in 3rd, 6th and 8th grade. The P(Leave) tabulations show the proportion of teachers with a given experience level who left Illinois public schools as of the end of the previous school year. Averages are shown within experience level for the pre-treatment period (1990-1993) and in the two school years immediately following the implementation of the early retirement incentive program (1994-1995). The next set of columns shows the distribution of experience among those exiting in the previous school year. Each of the first two columns sums to one in these estimates. The final set of tabulations shows the experience distribution of those entering the Illinois Public Schools in a given school year. Each of the first two columns sums to one in these tabulations.

**Table 3. OLS Estimates of the Effect of the Early Retirement Incentive Program on Teacher Composition and School Resources**

| Dependent Variable | Full Sample | 3rd Grade | 6th Grade | 8th Grade |
|---|---|---|---|---|
| Number Exiting in the Previous School Year | 0.078**<br>(0.013) | 0.081**<br>(0.011) | 0.085**<br>(0.024) | 0.075**<br>(0.019) |
| Dependent Variable Mean | 0.239 | 0.108 | 0.234 | 0.444 |
| Average Teacher Experience in Grade | -0.405**<br>(0.040) | -1.088**<br>(0.103) | -0.472**<br>(0.060) | -0.241**<br>(0.049) |
| Dependent Variable Mean | 15.52 | 15.39 | 15.25 | 16.09 |
| Number of New Teachers | 0.073**<br>(0.014) | 0.076**<br>(0.019) | 0.088**<br>(0.017) | 0.063**<br>(0.022) |
| Dependent Variable Mean | 0.368 | 0.156 | 0.405 | 0.635 |
| Pupil-Teacher Ratio | -0.032**<br>(0.015) | -0.076**<br>(0.031) | -0.045**<br>(0.020) | -0.014<br>(0.021) |
| Dependent Variable Mean | 19.84 | 19.88 | 19.92 | 19.66 |

Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1997 combined with school-level Illinois State Board of Education data on test scores.

Notes: The table shows estimates of $\beta_1$ from equation (1) in the text, using as dependent variables the number of teachers with 15 or more years of experience exiting in the previous year, average teacher experience in each grade, the number of new teachers and pupil-teacher ratios in each grade. Each cell presents results from a separate regression in the first column, and the estimates in each row in the final three columns come from one regression that includes interactions between $Post * Teachers \geq 15$ and *Post**Teachers* and grade indicators. Teachers who teach multiple grades are included in the estimates for each grade in which they teach. Teacher counts are based on all teachers within a grade and school in each year. All estimates include controls for the demographic variables included in Table 1, the total number of teachers in the grade interacted with a *Post* indicator, school fixed effects, and grade-by-year fixed effects. Standard errors clustered at the school-grade level are in parentheses: ** indicates statistical significance at the 5% level and * indicates statistical significance at the 10% level.

**Table 4. OLS Estimates of the Effect of the Early Retirement Incentive Program on the Number of Teachers of Difference Experience Levels**

| | Independent Variable: Post*Number $\geq 15$ | | | |
|---|---|---|---|---|
| Teacher Experience Level | Full Sample | 3rd Grade | 6th Grade | 8th Grade |
| 1 | 0.073**<br>(0.014) | 0.076**<br>(0.019) | 0.088**<br>(0.017) | 0.063**<br>(0.022) |
| Dep. Var. Mean | 0.368 | 0.156 | 0.405 | 0.635 |
| 2 | 0.060**<br>(0.014) | 0.059**<br>(0.018) | 0.081**<br>(0.024) | 0.047**<br>(0.019) |
| Dep. Var. Mean | 0.366 | 0.167 | 0.410 | 0.604 |
| 3 | 0.030**<br>(0.012) | 0.053**<br>(0.016) | 0.055**<br>(0.016) | 0.013<br>(0.019) |
| Dep. Var. Mean | 0.347 | 0.163 | 0.387 | 0.568 |
| 4 | 0.028**<br>(0.012) | 0.043**<br>(0.014) | 0.022**<br>(0.019) | 0.028**<br>(0.018) |
| Dep. Var. Mean | 0.317 | 0.159 | 0.336 | 0.528 |
| 5 | 0.022**<br>(0.012) | 0.025<br>(0.017) | -0.013<br>(0.015) | 0.038**<br>(0.019) |
| Dep. Var. Mean | 0.306 | 0.157 | 0.319 | 0.513 |
| 6-9 | 0.049**<br>(0.021) | 0.044*<br>(0.027) | 0.039<br>(0.030) | 0.053*<br>(0.032) |
| Dep. Var. Mean | 1.096 | 0.569 | 1.126 | 1.849 |
| 10-14 | 0.009<br>(0.022) | 0.068**<br>(0.028) | 0.016<br>(0.038) | -0.006<br>(0.032) |
| Dep. Var. Mean | 1.485 | 0.741 | 1.524 | 2.551 |
| 15-19 | -0.266**<br>(0.033) | -0.305**<br>(0.028) | -0.227**<br>(0.053) | -0.280**<br>(0.047) |
| Dep. Var. Mean | 1.828 | 0.880 | 1.847 | 3.235 |
| 20-24 | -0.178**<br>(0.032) | -0.048*<br>(0.027) | -0.216**<br>(0.056) | -0.178**<br>(0.044) |
| Dep. Var. Mean | 1.826 | 0.763 | 1.809 | 3.452 |
| 25-29 | -0.105**<br>(0.022) | -0.011<br>(0.021) | -0.099**<br>(0.035) | -0.125**<br>(0.031) |
| Dep. Var. Mean | 1.048 | 0.453 | 1.007 | 2.004 |
| 30-34 | -0.010<br>(0.015) | 0.009<br>(0.013) | -0.023<br>(0.019) | -0.008<br>(0.023) |
| Dep. Var. Mean | 0.512 | 0.190 | 0.504 | 1.007 |
| 35-39 | 0.001<br>(0.009) | 0.007<br>(0.007) | 0.012<br>(0.016) | -0.006<br>(0.012) |

| | | | | |
|---|---|---|---|---|
| Dep. Var. Mean | 0.117 | 0.048 | 0.105 | 0.238 |
| 40+ | -0.002<br>(0.002) | 0.0002<br>(0.003) | -0.0003<br>(0.0004) | -0.003<br>(0.002) |
| Dep. Var. Mean | 0.017 | 0.011 | 0.017 | 0.024 |

Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1997 combined with school-level Illinois State Board of Education data on test scores.

Notes: The table shows estimates of $\beta_1$ from equation (1) in the text, using as dependent variables the number of teachers with a given level of experience in each grade and year. Each cell presents results from a separate regression in the first column, and the estimates in each row in the final three columns come from one regression that includes interactions between $Post * Teachers \geq 15$ and *Post***Teachers* and grade indicators. Teachers who teach multiple grades are included in the estimates for each grade in which they teach. Teacher counts are based on all teachers within a grade and school in each year. All estimates include controls for the demographic variables included in Table 1, the total number of teachers in the grade interacted with a *Post* indicator, school fixed effects, and grade-by-year fixed effects. Standard errors clustered at the school-grade level are in parentheses: ** indicates statistical significance at the 5% level and * indicates statistical significance at the 10% level.

**Table 5. OLS Estimates of the Effect of the Early Retirement Incentive Program on Student Test Scores**

| | All Teachers | | Subject-Specific Teachers | |
|---|---|---|---|---|
| Ind. Var. | Math | Reading | Math | Reading |
| Post*Number $\geq 15$ | 0.003<br>(0.004) | 0.009**<br>(0.003) | 0.013*<br>(0.008) | 0.013*<br>(0.007) |
| Post*Total Teachers | 0.0001<br>(0.002) | -0.004*<br>(0.002) | -0.004<br>(0.006) | -0.005<br>(0.005) |

Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1997 combined with school-level Illinois State Board of Education data on test scores.
Notes: Each column presents results from a separate regression. Teachers who teach multiple grades are included in the estimates for each grade in which they teach. Teachers who teach in self-contained classrooms are assumed to teach both math and English. All estimates include controls for the demographic variables included in Table 1, school fixed effects, and grade-by-year fixed effects. Standard errors clustered at the school-grade level are in parentheses: ** indicates statistical significance at the 5% level and * indicates statistical significance at the 10% level.

**Table 6. OLS Estimates of the Effect of the Early Retirement Incentive Program on Student Test Scores, by Grade**

| | All Teachers | | Subject-Specific Teachers | |
|---|---|---|---|---|
| Ind. Var. | Math | Reading | Math | Reading |
| Post*Number $\geq 15$ *I(Grade=3) | 0.002<br>(0.010) | -0.009<br>(0.008) | 0.010<br>(0.013) | -0.003<br>(0.010) |
| Post*Number $\geq 15$ *I(Grade=6) | -0.00002<br>(0.005) | 0.006<br>(0.004) | 0.004<br>(0.010) | 0.016*<br>(0.009) |
| Post*Number $\geq 15$ *I(Grade=8) | 0.005<br>(0.005) | 0.013**<br>(0.005) | 0.032*<br>(0.018) | 0.030*<br>(0.017) |

Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1997 combined with school-level Illinois State Board of Education data on test scores.
Notes: Each column presents results from a separate regression. Teachers who teach multiple grades are included in the estimates for each grade in which they teach. Teachers who teach in self-contained classrooms are assumed to teach both math and English. All estimates include controls for the demographic variables included in Table 1, the total number of teachers in a grade (or the total number of subject-specific teachers in the grade in the second two columns) interacted with a *Post* indicator, school fixed effects, and grade-by-year fixed effects. Standard errors clustered at the school-grade level are in parentheses: ** indicates statistical significance at the 5% level and * indicates statistical significance at the 10% level.

**Table 7. OLS Estimates of the Effect of the Early Retirement Incentive Program on Student Test Scores, by Income, Race and Baseline Scores**

| | Post*Number $\geq 15$ | |
|---|---|---|
| School Type | Math | Reading |
| Lowest Income | -0.001 | 0.014 |
| Schools | (0.008) | (0.008) |
| All Other | 0.002 | 0.009** |
| Schools | (0.004) | (0.004) |
| | | |
| Lowest % | 0.004 | 0.018** |
| White Schools | (0.007) | (0.007) |
| All Other | 0.004 | 0.006 |
| Schools | (0.004) | (0.004) |
| | | |
| Lowest Baseline | 0.010 | 0.019** |
| Score Schools | (0.008) | (0.008) |
| All Other | 0.003 | 0.008** |
| Schools | (0.004) | (0.004) |

Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1997 combined with school-level Illinois State Board of Education data on test scores.
Notes: Each cell presents results from a separate regression. Teachers who teach multiple grades are included in the estimates for each grade in which they teach. The lowest income, percent white and baseline score schools are the bottom 25% of schools on each measure, measured in the pre-treatment period. All estimates include controls for the demographic variables included in Table 1 (except for the variable on which the sample is cut), the total number of teachers interacted with a *Post* indicator, school fixed effects, and grade-by-year fixed effects. Standard errors clustered at the school-grade level are in parentheses: ** indicates statistical significance at the 5% level and * indicates statistical significance at the 10% level.

**Table 8. OLS Estimates of the Effect of the Early Retirement Incentive Program on Teacher Composition and School Resources, by Income, Race and Baseline Scores**

| Dependent Variable | Lowest Income | Other Income | Lowest % White | Other % White | Lowest Math | Other Math | Lowest Reading | Other Reading |
|---|---|---|---|---|---|---|---|---|
| Number Exiting in Previous Year | 0.020 | 0.085** | 0.049* | 0.091** | 0.038 | 0.087** | 0.046* | 0.086** |
| | (0.034) | (0.016) | (0.026) | (0.016) | (0.029) | (0.016) | (0.028) | (0.016) |
| Dep. Var. Mean | 0.177 | 0.253 | 0.264 | 0.231 | 0.230 | 0.241 | 0.224 | 0.243 |
| Average Teacher Experience | -0.209** | -0.465** | -0.203** | -0.496** | -0.214** | -0.468** | -0.218** | -0.473** |
| | (0.083) | (0.037) | (0.081) | (0.040) | (0.098) | (0.037) | (0.096) | (0.038) |
| Dep. Var. Mean | 15.707 | 15.469 | 15.612 | 15.480 | 15.717 | 15.453 | 15.589 | 15.490 |
| Number of New Teachers | -0.060** | 0.104** | -0.024 | 0.107** | -0.062** | 0.107** | -0.008 | 0.108** |
| | (0.025) | (0.013) | (0.027) | (0.014) | (0.027) | (0.013) | (0.005) | (0.013) |
| Dep. Var. Mean | 0.248 | 0.395 | 0.342 | 0.377 | 0.312 | 0.384 | 0.314 | 0.384 |
| Pupil-Teacher Ratio | 0.020 | -0.046** | 0.004 | -0.053** | 0.018 | -0.048** | 0.027 | -0.050** |
| | (0.029) | (0.014) | (0.028) | (0.014) | (0.036) | (0.013) | (0.036) | (0.013) |
| Dep. Var. Mean | 20.036 | 19.798 | 19.852 | 19.837 | 20.206 | 19.736 | 20.254 | 19.718 |

Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1997 combined with school-level Illinois State Board of Education data on test scores.

Notes: The table shows estimates of $\beta_1$ from equation (1) in the text, using as dependent variables the number of teachers with 15 or more years of experience exiting in the previous year, average teacher experience in each grade, the number of new teachers and pupil-teacher ratios in each grade. Each cell presents results from a separate regression. Teachers who teach multiple grades are included in the estimates for each grade in which they teach. The lowest income, percent white and baseline score schools are the bottom 25% of schools on each measure, measured in the pre-treatment period. All estimates include controls for the demographic variables included in Table 1 (except for the variable on which the sample is cut), the total number of teachers interacted with a *Post* indicator, school fixed effects, and grade-by-year fixed effects. Standard errors clustered at the school-grade level are in parentheses: ** indicates statistical significance at the 5% level and * indicates statistical significance at the 10% level.

**Table 9. OLS Estimates of the Effect of the Early Retirement Incentive Program on Student Test Scores – Including Other Grades**

| | All Teachers – Excluding 8th Grade | | All Teachers – Including Higher Grade Teachers | |
|---|---|---|---|---|
| Ind. Var. | Math | Reading | Math | Reading |
| Post*Number $\geq 15$ | 0.001<br>(0.005) | 0.002<br>(0.004) | 0.001<br>(0.005) | 0.004<br>(0.004) |
| Post*Total Teachers | 0.001<br>(0.003) | 0.0004<br>(0.003) | 0.002<br>(0.003) | -0.00002<br>(0.003) |
| Post*Number Other $\geq 15$ | | | -0.005<br>(0.005) | -0.006<br>(0.004) |
| Post*Total Other Teachers | | | -0.001<br>(0.003) | 0.002<br>(0.003) |

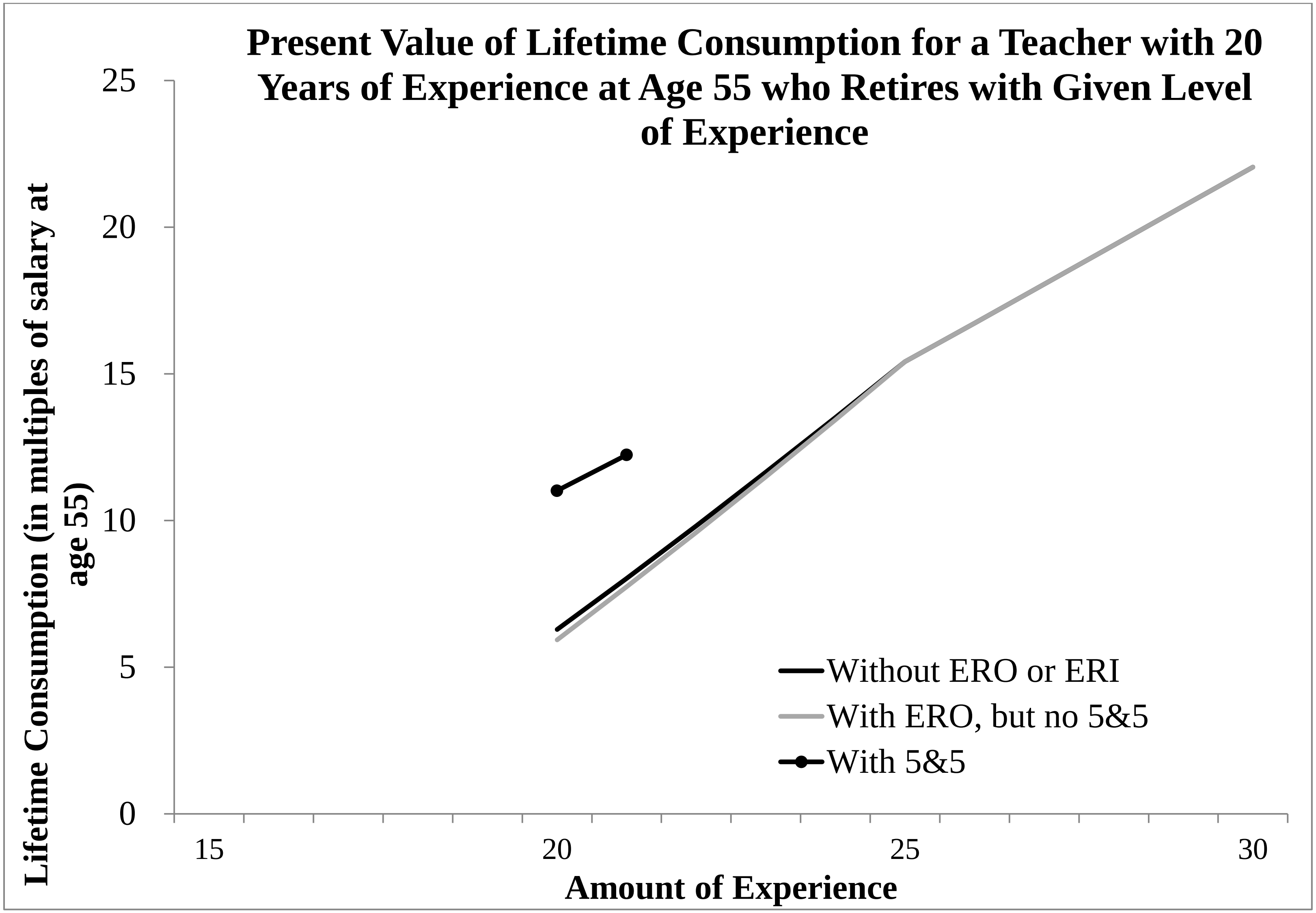
Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1997 combined with school-level Illinois State Board of Education data on test scores.
Notes: Each column presents results from a separate regression. Eighth grades are excluded from the estimates. Teachers who teach multiple grades are included in the estimates for each grade in which they teach. All estimates include controls for the demographic variables included in Table 1, school fixed effects, and grade-by-year fixed effects. "Other" refers to all teachers in the school in higher grades. Standard errors clustered at the school-grade level are in parentheses: ** indicates statistical significance at the 5% level and * indicates statistical significance at the 10% level.

**Table 10. Robustness Checks**

| | All Teachers | | Subject-Specific Teachers | |
|---|---|---|---|---|
| Specification | Math | Reading | Math | Reading |
| Post* Number $\geq 20$ | 0.002<br>(0.003) | 0.009**<br>(0.003) | 0.008<br>(0.008) | 0.008<br>(0.007) |
| Post*Percentage $\geq 15$ | 0.030<br>(0.039) | 0.063*<br>(0.035) | 0.007<br>(0.019) | 0.007<br>(0.019) |
| Post* Number $\geq 15$ (Excluding Chicago) | 0.004<br>(0.004) | 0.009**<br>(0.003) | 0.015*<br>(0.007) | 0.013*<br>(0.007) |
| Post* Number $\geq 15$ (Excluding Missing) | 0.002<br>(0.004) | 0.009**<br>(0.003) | 0.013*<br>(0.008) | 0.013*<br>(0.007) |

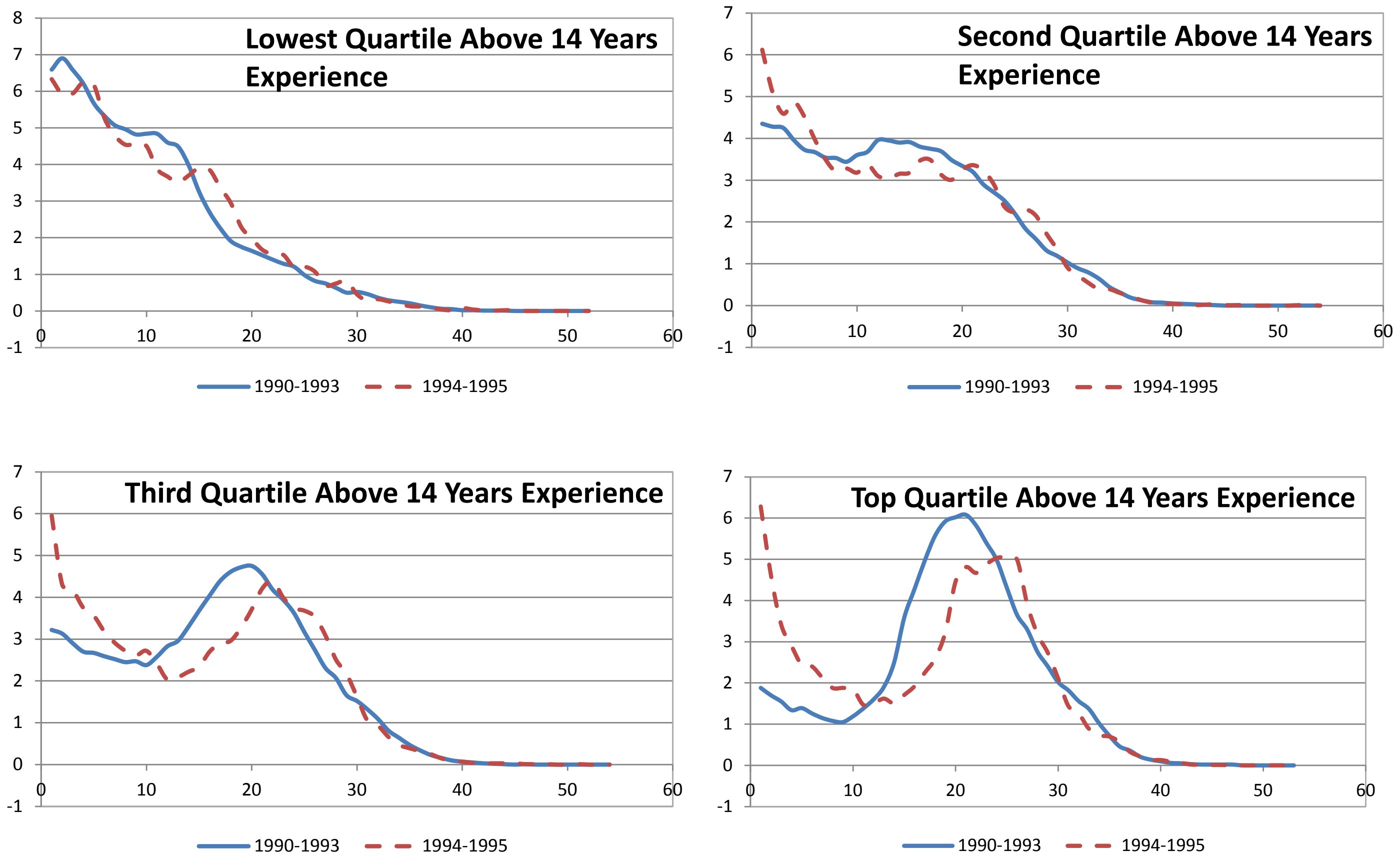
Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1997 combined with school-level Illinois State Board of Education data on test scores.
Notes: Each column presents results from a separate regression. Teachers who teach multiple grades are included in the estimates for each grade in which they teach. Teachers who teach in self-contained classrooms are assumed to teach both math and English. All estimates include controls for the demographic variables included in Table 1, school fixed effects, and grade-by-year fixed effects. The estimates in the first, third and fourth rows include an interaction between the total number of teachers in the grade (or the total number of subject-specific teachers) and a *Post* indicator. Standard errors clustered at the school-grade level are in parentheses: ** indicates statistical significance at the 5% level and * indicates statistical significance at the 10% level.

**Figure 1. Budget Constraints for a Representative Teacher in IPS, With and Without ERI (5+5)**




Notes: Based on authors' calculations. The vertical axis measures the present value of consumption for a teacher who is age 55 and has 20 years of experience when the ERI becomes available. The vertical axis presents the present discounted value of lifetime consumption associated with retirement with a given amount of experience (without taking time off) and under a certain retirement plan. Consumption in each year is defined as income (salary if working and retirement benefit if retired) minus fees (if retirement incentive is purchased). Salary and benefits are expected to grow by 3 percent per year and the discount rate is 0.97. The present value of consumption is measured at age 55 and is presented in multiples of salary at age 55.

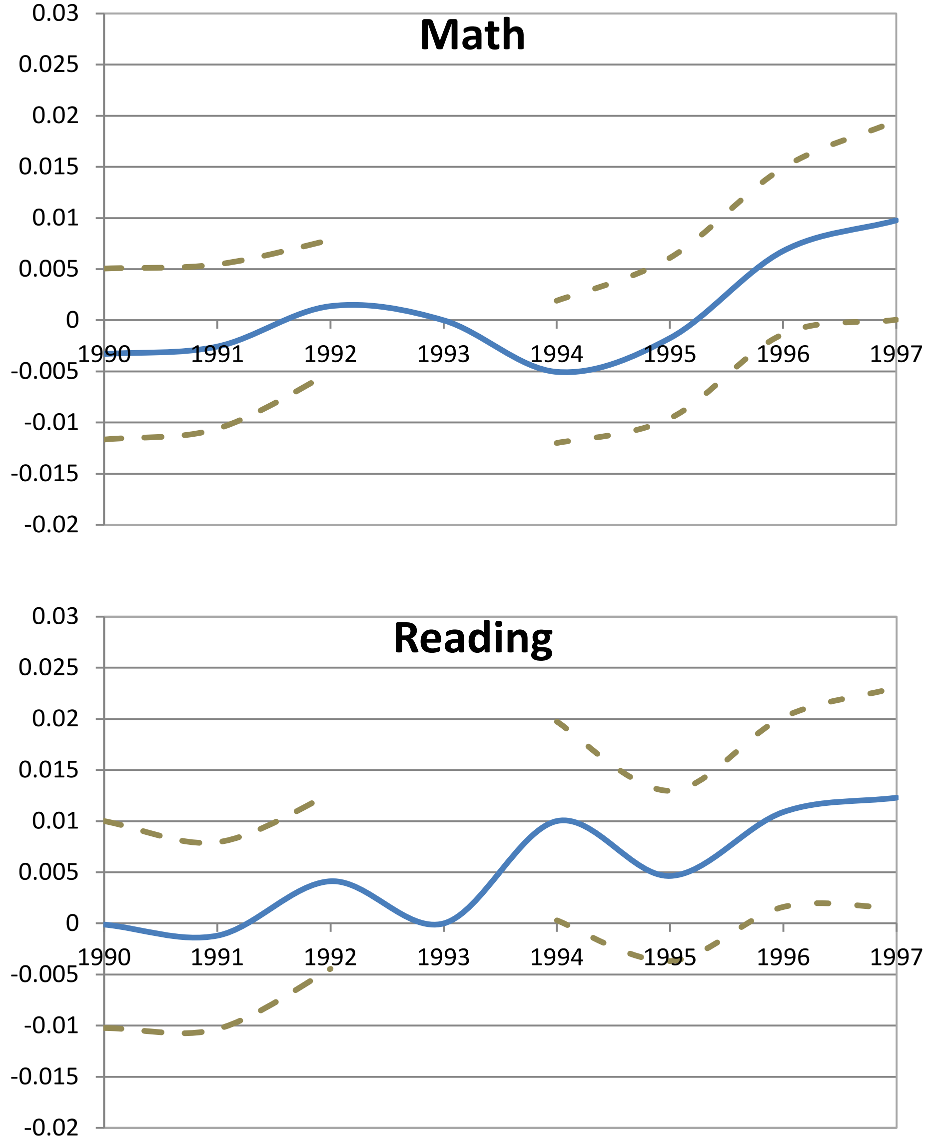
**Figure 2. Distribution of Teacher Experience Pre- and Post-Teacher Early Retirement Incentive Program, by Quartiles of the Percent of Grade-Specific Teachers with 15 or More Years of Experience Prior to 1994**




Source: Illinois Teacher Service Record data from 1990 through 1995 combined with school-level Illinois State Board of Education data on test scores.

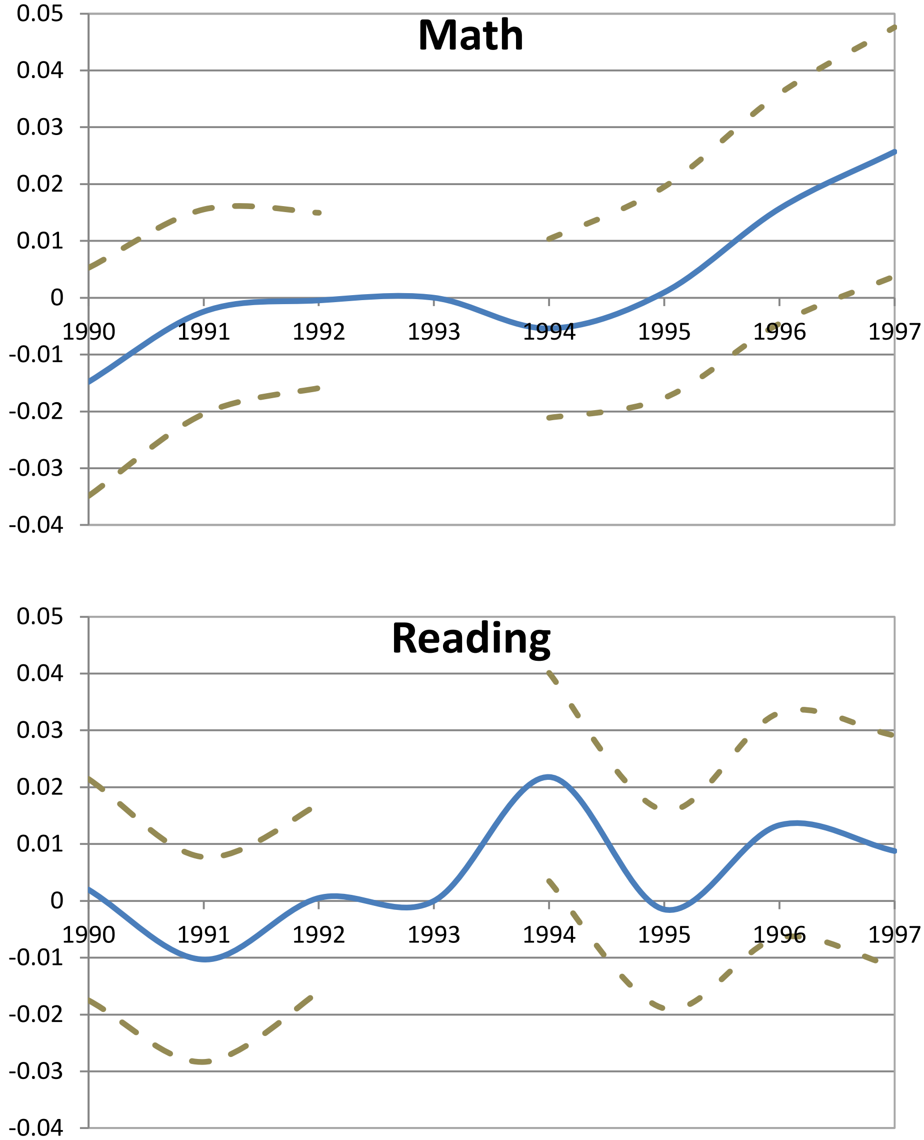
Notes: All years refer to the calendar year in which a school year ends, and distributions are based on all teachers in 3rd, 6th and 8th grade. Each panel of the figure shows experience distributions for the pre-treatment period (1990-1993) and the period immediately following the early retirement incentive program implementation (1994-1995) by the quartile of the percent of teachers with 20 or more years of experience in the pre-treatment period.

**Figure 3. Event Study Estimates of the Effect of the Early Retirement Incentive Program on Student Test Scores, Using Grade-Specific Teacher Counts**



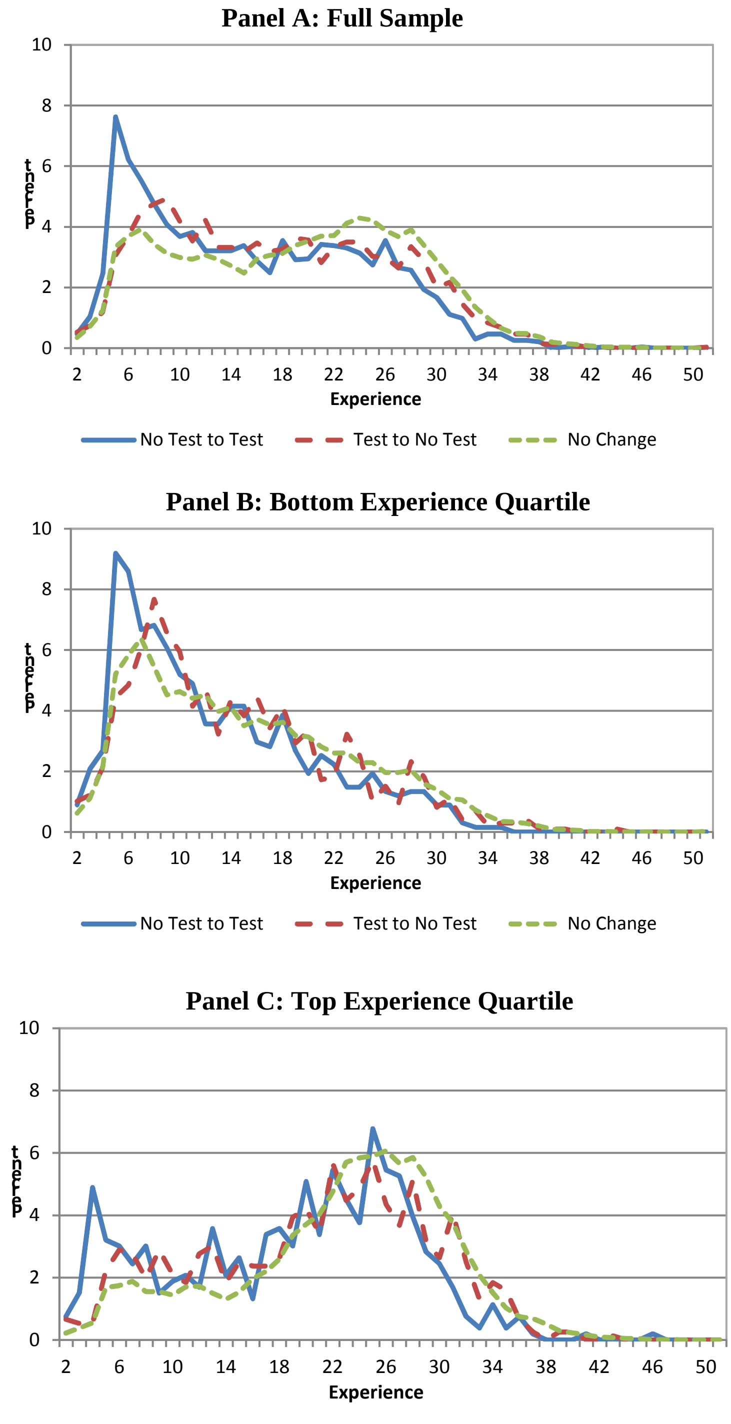

Notes: Estimates and 95% confidence intervals from interactions of year fixed effects and the number of teachers with 15 or more years of experience. Years are indexed by the calendar year in which a school year ends. The 1993 coefficient is set to zero, so there is no standard error bound for this year. Estimates include grade-by-year and school fixed effects, the demographic characteristics shown in Table 1, and interactions between year fixed effects and the number of teachers in each grade pre-ERI.

**Figure 4. Event Study Estimates of the Effect of the Early Retirement Incentive Program on Student Reading Test Scores, Using Subject- and Grade-Specific Teacher Counts**




Notes: Estimates and 95% confidence intervals from interactions of year fixed effects and the number of subject-specific teachers with 15 or more years of experience. Years are indexed by the calendar year in which a school year ends. The 1993 coefficient is set to zero, so there is no standard error bound for this year. Estimates include grade-by-year and school fixed effects, the demographic characteristics shown in Table 1, and interactions between year fixed effects and the number of subject-specific teachers in each grade pre-ERI.

**Figure 5. Pre-ERI Experience Distributions among Teachers Who Were in IPS Pre- and Post-ERI, by Whether They Shifted Across Tested and Non-Tested Grades**




Notes: Sample include only those teachers who taught in IPS both pre- and post-ERI. "Test grades" are 3, 6, and 8, and all other grades below 8 are "no test" grades. Top experience and bottom experience refer to quartiles of the pre-ERI percent of teachers with 15 or more years of experience.

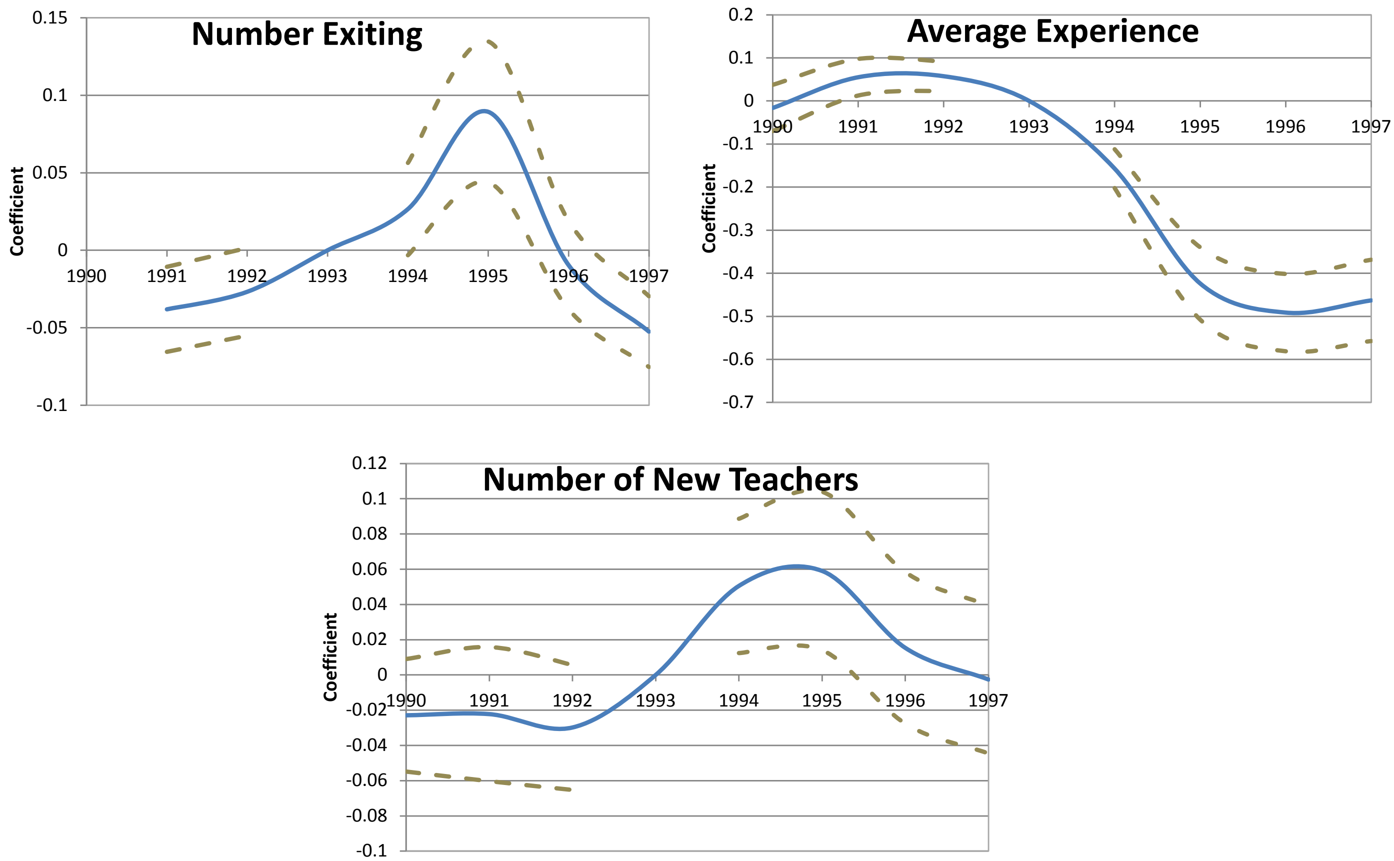
**Table A-1. OLS Estimates of the Effect of the Early Retirement Incentive Program on the Number of Teachers of Difference Experience Levels in Disadvantaged Schools**

| | Independent Variable: Post*Number $\geq 15$ | | | |
|---|---|---|---|---|
| Teacher Experience Level | Low Income | Low White | Low Math | Low Reading |
| 1 | -0.060**<br>(0.025) | -0.024<br>(0.027) | -0.062**<br>(0.027) | -0.008<br>(0.005) |
| 2 | -0.009<br>(0.027) | 0.024<br>(0.030) | -0.006<br>(0.032) | 0.002<br>(0.031) |
| 3 | 0.015<br>(0.023) | -0.005<br>(0.023) | -0.001<br>(0.022) | -0.003<br>(0.021) |
| 4 | 0.012<br>(0.028) | 0.025<br>(0.027) | 0.013<br>(0.023) | 0.011<br>(0.023) |
| 5 | 0.032<br>(0.027) | 0.064**<br>(0.020) | 0.056**<br>(0.024) | 0.057**<br>(0.023) |
| 6-9 | 0.014<br>(0.049) | 0.032<br>(0.040) | 0.027<br>(0.043) | 0.021<br>(0.042) |
| 10-14 | -0.024<br>(0.043) | 0.014<br>(0.035) | 0.001<br>(0.043) | -0.015<br>(0.046) |
| 15-19 | -0.291**<br>(0.051) | -0.269**<br>(0.054) | -0.268**<br>(0.053) | -0.267**<br>(0.052) |
| 20-24 | -0.023<br>(0.070) | -0.104*<br>(0.056) | -0.104<br>(0.071) | -0.114*<br>(0.070) |
| 25-29 | -0.121**<br>(0.052) | -0.137**<br>(0.048) | -0.008**<br>(0.003) | -0.174**<br>(0.056) |
| 30-34 | 0.006<br>(0.024) | 0.007<br>(0.029) | -0.163**<br>(0.056) | 0.010<br>(0.033) |
| 35-39 | 0.003<br>(0.019) | -0.004<br>(0.014) | 0.001<br>(0.033) | 0.0004<br>(0.018) |
| 40+ | -0.005<br>(0.004) | -0.002<br>(0.005) | 0.012<br>(0.018) | -0.007<br>(0.005) |

Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1997 combined with school-level Illinois State Board of Education data on test scores.

Notes: The table shows estimates of $\beta_1$ from equation (1) in the text, using as dependent variables the number of teachers with a given level of experience in each grade and year. Each cell presents results from a separate regression. Teachers who teach multiple grades are included in the estimates for each grade in which they teach. The lowest income, percent white and baseline score schools are the bottom 25% of schools on each measure, measured in the pre-treatment period. All estimates include controls for the demographic variables included in Table 1 (except for the variable on which the sample is cut), the total number of teachers interacted with a *Post* indicator, school fixed effects, and grade-by-year fixed effects. Standard errors clustered at the school-grade level are in parentheses: ** indicates statistical significance at the 5% level and * indicates statistical significance at the 10% level.

**Figure A-1. Event Study Estimates of the Effect of the Early Retirement Incentive Program on Teacher Composition**




Notes: Estimates and 95% confidence intervals from interactions of year fixed effects and the number of teachers with 15 or more years of experience. Years are indexed by the calendar year in which a school year ends. The 1993 coefficient is set to zero, so there is no standard error bound for this year. Estimates include grade-by-year and school fixed effects, the demographic characteristics shown in Table 1, and interactions between year fixed effects and the number of teachers in each grade.