



Working Papers

www.cesifo.org/wp

Early Retirement Incentives and Student Achievement

Maria D. Fitzpatrick
Michael F. Lovenheim

CESIFO WORKING PAPER NO. 4347
CATEGORY 5: ECONOMICS OF EDUCATION
JULY 2013

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

Early Retirement Incentives and Student Achievement

Abstract

Early retirement incentives (ERIs) are increasingly prevalent in education as districts seek to close budget gaps by replacing expensive experienced teachers with lower-cost newer teachers. Combined with the aging of the teacher workforce, these ERIs are likely to change the composition of teachers dramatically in the coming years. We use exogenous variation from an ERI program in Illinois in the mid-1990s to provide the first evidence in the literature of the effects of large-scale teacher retirements on student achievement. We find the program did not reduce test scores; likely, it increased them, with positive effects most pronounced in lower-SES schools.

JEL-Code: J260, I210, I280, H750.

Keywords: teacher retirement, early retirement incentives, teacher experience.

Maria D. Fitzpatrick
Department of Policy Analysis and
Management / Cornell University
103 Martha Van Rensselaer Hall
USA – Ithaca, NY 14853
maria.d.fitzpatrick@cornell.edu

Michael F. Lovenheim
Department of Policy Analysis and
Management / Cornell University
103 Martha Van Rensselaer Hall
USA – Ithaca, NY 14853
mfl55@cornell.edu

July 2013

We would like to thank Steve Rivkin, Jonah Rockoff and seminar participants at Cornell University, CESifo, North Carolina State University, the NBER Economics of Education Working Group and the American Economic Association Annual Meetings for helpful comments and suggestions. Funding for Fitzpatrick from the National Institute on Aging, through Grant Number T32-AG000186 to the National Bureau of Economic Research, is gratefully acknowledged. All errors and omissions are our own.

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 ERI 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 influence 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, when many teachers become eligible for full-retirement benefits based on their age and experience, in the coming decade we can expect a large proportion of teachers to be on the margin where an ERI may be particularly attractive.² Indeed, in the last couple of years, many states already have seen an exodus of teachers into retirement as retirement incentives have been offered and 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).

The growing literature on teacher retirement focuses largely on the role of the retirement incentives embedded in pension plans (Ferguson, Strauss and Vogt 2006; Costrell and Podgursky 2009; Costrell and McGee 2010; Brown and Laschever 2012; Brown 2013) 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.

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 Ferguson, 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)

Ex ante, it is unclear what the effects of large-scale teacher retirements, such as those resulting from an ERI, 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. The evidence of the strong relationship between experience and effectiveness in the classroom (Wiswall 2013; Rivkin, Hanushek and Kain 2005; Rockoff 2004) suggests teacher retirements could reduce student achievement. Even among teachers who have the same amount of experience, teacher quality varies substantially (Goldhaber and Hansen 2010). If teachers with better job opportunities outside of Illinois Public Schools (IPS) are the most likely to retire, and if wages outside teaching are positively correlated with teacher quality (Chingos and West 2012), then the offer of an ERI would negatively affect student test scores.

On the other hand, teachers who are near retirement 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 who desire to retire early. Alternatively, if productivity is negatively correlated with disutility from teaching, the teachers who choose to take up the ERI may be those that are least productive. Family and personal circumstances also influence the labor-leisure decision in ways that lead to ambiguous predictions of the effect of ERIs on achievement. Finally, principals and administrators may respond to large losses of experienced teachers, e.g. by decreasing class sizes, changing the assignment of teachers to students or purchasing additional non-teacher resources.

Despite the increasing importance of understanding how student academic achievement responds to teacher retirement behavior, previous research has not been able to credibly estimate this relationship. 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.

We use a natural experiment brought about by a two-year ERI program in Illinois to identify the effect of ERIs 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 Specifically, we exploit a temporary ERI that was offered to all public school

⁴ Mahler (2012) and McGee (2012) examine how the responsiveness of teachers' retirement to the nonlinearities in their pension benefit calculation formulas 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 pension incentives. Responses to the fixed long-standing benefit system may be different than responses to a short-lived incentive offer like the ERI, however.

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 1992-1993 and 1993-1994 school years, the state offered an ERI, 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 before the 1993-1994 and 1994-1995 school years. 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 the 1989-1990 to 1996-1997 school years on 3rd, 6th and 8th grade state standardized math 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.

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 at a school 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 2013), 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.

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 (Lankford, Loeb and Wyckoff 2002; Feng et al. 2012). We explore the role of heterogeneity in responses across

schools by school-average income, percent white and by baseline test scores. There is 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. Although the differences are not statistically significant, to the extent that there is a difference, we show that some of it is driven by how schools were affected by the ERI program. More disadvantaged schools experienced larger total declines in the size of the teacher workforce and were more likely to replace retirees with teachers with some prior experience (rather than novices).

It is important to consider what these results contribute to our understanding of the returns to experience and to the effects of the impending swath of teacher retirements as the Baby Boomers age. 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 has failed to reach a consensus on whether there are returns to experience for middle- and end-of-career teachers.⁵ Our results support the argument that teachers who are very high in the experience distribution, or who are close to retirement, are less productive. However, 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 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 ERI 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, the employee contribution rate was

⁵ For example, see Papay and Kraft (2011), Wiswall (2013), Clotfelter et al. (2006), Rivkin, Hanushek and Kain (2005), and Rockoff (2004).

9 percent of earnings, and 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.⁶

In general, retirement benefits can be claimed by members of the TRS when they terminate active service with 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 with 5 years of service.⁷ 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.⁸

In 1992-1993 and 1993-1994, employees were offered an 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.⁹ The fee for employees was 4 percent of the highest annual salary for the past five years for each of the five additional years of age and service purchased; the fee for employers was 12 percent of the employee's highest annual salary from the last five years for each additional year purchased.¹⁰

Chicago has a separate, but very similar, pension system for its teachers, called the Chicago Teachers Pension Fund (CTPF). In terms of benefit calculation and the offer of the ERI, CTPF works exactly like TRS. Moreover, the TRS and CTPF have an agreement of

⁶ 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.

⁷ 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.

⁸ 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.

⁹ A third window at the end of the 1994-1995 school year was allowed for the small number of retirements delayed by employers.

¹⁰ 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 did not prohibit retirement.

complete reciprocity, whereby an employee who has spent time working in both systems can have the creditable service in one transferred to the other. The City of Chicago had to separately negotiate its own 5+5 agreement with the Chicago Teachers' Union, but it did so in time for the ERI program to take place at the same time across the state. Thus, teachers across the entire state of Illinois were subject to the same ERI eligibility and benefit rules.

Lifetime budget constraints based on observed teacher salaries and the benefits formulas with and without the ERO and with and without ERI are illustrated in Figure 1.A. 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 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). Note that this does not include any potential income from other sources, such as a job outside of IPS after retiring from teaching. 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, both with and without the ERO, are indicated by the solid gray and black lines, 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, this discount is no longer taken since she has more than 10 years of experience, which causes the kink in her budget constraint.

More relevant for our setting is the budget segment with two circular markers, representing the budget constraint under the ERI program. 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 1992-1993 or 1993-1994 school years, which is why the change is represented as a budget segment available to her only if she retires with 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 and consumption, 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, as shown in Figure 1.B, the ERI represented large potential gains in lifetime consumption for most teachers who were at least 50 years old and who had at least 15 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. The Teacher Service Record (TSR) is an administrative dataset collecting information about employees of the 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 from the 1989-1990 to 1996-1997 school years, because the earliest available data are from the 1989-1990 school year, and in 1998 the Illinois legislature changed the teacher benefit formula in ways that could influence teacher retirement decisions.¹¹ Note that 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. Below, we describe each of the data sets we use 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 1990 and 1997. In what follows, we detail several choices regarding sample restriction and variable definitions. We test the sensitivity of our results to these choices in Section 5.5.

Standardized testing in Illinois is focused on 3rd, 6th and 8th 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 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

¹¹ See Fitzpatrick (2013) for a detailed description of the 1998 policy change as well as the extent to which teachers value these benefits.

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.¹²

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.¹³ Using the empirical distribution of teachers in the 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.¹⁴ 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.

¹² 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.

¹³ 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.

¹⁴ Including experience outside of Illinois adds little variation to overall experience, as total experience and experience in Illinois have a correlation of 0.98.

Table 1 provides descriptive statistics, overall and by grade, for this variable.¹⁵ 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 about 56 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 experience composition of teachers across schools prior to the ERI program.

We also use the teacher-level data to calculate exit rates of experienced teachers,¹⁶ 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 observe teacher retirement or ERI program takeup directly. 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 one 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.

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/state assistance. Variable means and standard deviations, as well as school counts, are included in Table 1.

4 **Empirical Methodology**

¹⁵ 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.

¹⁶ We focus on “experienced” teacher exits, as this is one of the treatments of interests. All results and conclusions regarding exits are unchanged if we use total exits instead. These results are available upon request.

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. Recall from the discussion in Section 2 that a teacher's eligibility to collect a pension benefit is completely determined by her age and experience. The earliest age at which retirement benefit collection could have taken place before the ERI is 55; with the ERI, teachers can now retire at age 50 or older. Most teachers aged 50 years or older have at least 15 years of experience, so we expect the ERI will affect the retirement behavior of teachers with at least 15 years of experience disproportionately.

The prediction that teachers with at least 15 years of experience were the most likely to retire is borne out in the data. In Panel A of Table 2, we show exit rates in the pre-treatment years (1990-1993, in Column (1)), in the treatment years (1994-1995, in Column (2)), and in the post-treatment years (1996-1997, in Column (3)) by experience level.¹⁷ In the subsequent columns, we report the differences in these experience-level-specific exit rates between the treatment and pre-treatment periods (Column (4)) and the post-treatment and pre-treatment periods (Column (5)).

The differences in the fourth 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 and 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.

As shown in Column (5), the post-treatment exit rates of teachers return to their pre-treatment levels or, for the most senior teachers, fall below their pre-treatment levels. For example, among teachers with 25 or more years of experience exit rates in the post-treatment period are less than half as large as exit rates in the pre-treatment period. As we would expect, the teachers that remain after the ERI-offer-period are those with stronger labor force attachment.

Information in Panel B of Table 2 shows how the ERI program affected the distribution of experience among teachers exiting. Consistent with the fact that this program increased

¹⁷ Note that the exit and experience tabulations in Table 2 show the proportion of teachers with a given experience level leaving and the distribution of experience among those exiting in the previous school year. Thus, 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.

retirement among experienced teachers, there was a marked outward shift in the experience of exiting teachers.

In calculations not reported in Table 2, we also examine the experience distribution of entering teachers before, during and after the ERI program period. The largest change induced by the ERI program was a 10 percentage point (or a 19.6%) increase in the proportion of new teachers (those with one year or less of experience) among those entering IPS. 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.

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) by quartile of the proportion of teachers in grades 3, 6 and 8 with experience greater than or equal to 15 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:

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

where Y_{igt}^s is the standardized test score in grade g for subject s in school i and year t . The variable $Teachers \geq 15$ is the average number of teachers in a given grade with at least 15 years of experience pre-1994. $Teachers$ is the average total number of teachers in a grade and school in the pre-ERI period and $Post$ is an indicator variable equal to 1 for school years after 1993. The vector X contains the set of school-by-year demographic variables discussed in Section 3 as well as a quadratic in school-grade-year enrollment, and the model includes both school-by-grade (

δ_{ig}) and grade-by-year (φ_{ig}) fixed effects.¹⁸ 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 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 β_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 school-grades with fewer or more potentially affected teachers, as measured prior to ERI introduction. 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 endogenous responses to the treatment itself.

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.¹⁹ 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 β_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 ERI 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

¹⁸ 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.

¹⁹ 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.

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 in the school-grade, and school-grade 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 in turn, 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* ≥ 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). Having one more teacher with 15 or more years of experience pre-ERI increased the number of experienced teachers exiting in the ERI period by 0.08 overall, with estimates that vary little across grades. Relative to the baseline, this represents a 32.6% increase.

The increased exit among more experienced teachers was accompanied by declines in average teacher experience of 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 part by an increase in 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 effect is largest in levels in grade 6, at 0.088, but relative to baseline it is largest in 3rd grade, at 48.7% . 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 when 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 slightly in schools with more experienced teachers due to this program. This average reduction is driven entirely by the 6th and 8th grades, which shrink by -0.265 and -0.364 teachers, respectively, for each teacher with more than 14 years of experience pre-ERI. The 3rd grade teacher workforce is basically unchanged, with a slight increase of 0.02 teachers for each pre-ERI experienced teacher. None of these results represent sizable changes in the number of teachers.

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 this type of behavior. 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 insignificant effect of -0.009. The pupil-teacher ratio effects are largest in absolute value in 3rd grade, at 0.183, but none of the estimates is statistically significant at conventional levels. Even the 3rd grade estimate is very small relative to baseline (1 percent), much smaller than the decreases in class sizes that have shown positive effects in the literature.²⁰

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 from models similar to equation (1), except the *Post*Teachers* ≥ 15 and *Post*Teachers* have been replaced with year indicators interacted with *Teachers* ≥ 15 and *Teachers*. The figures show the *Teachers* $\geq 15 * I(\text{year})$ 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

²⁰ 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.

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.²¹ 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. Overall, the estimates in Figures 3 and 4 support our identification strategy and suggest that the changes to schools shown in Tables 3 and 4 brought about by ERI-induced retirements did not reduce student test scores.

Due to the demanding nature of the event study 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. 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.

²¹ Appendix Figure A-1 contains event study estimates for experienced 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.

One informative way to scale these results is per exiting teacher. By dividing the estimates in Table 5 by the experienced teacher exit estimates in Table 3, one can calculate such an instrumental variables estimate. We caution, however, that the exclusion restriction is unlikely to be met; as indicated in Table 3, the ERI program affected the teacher composition of schools along a number of dimensions. Attributing all of the test score effect to any one of these composition outcomes would be incorrect, but it still is helpful to scale these results in terms of each exiting teacher. Performing this calculation for estimates in the first two columns of Table 4 yields IV estimates of 0.038 and 0.115 for math and reading, respectively. These are school-grade-level standard deviations, however. The variance of school-grade test scores is approximately 15% of the student-level variance in test scores, so one can convert to student-level standard deviation by dividing by 2.6 (the square root of 0.15). Thus, for each teacher who leaves under the ERI, our estimates indicate test scores increase by 0.01 student-level standard deviations for math and 0.04 student-level standard deviations for reading. While small in magnitude, we show in Section 6 that these test score increases are slightly more cost-effective than comparable effects from class size increases.

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 8th grade in both subjects that are significant at the 10% level. This larger effect could be due to the fact that each 8th grade teacher teaches a larger proportion of the students, which would increase the impact of any one teacher retiring. 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 median retiring teacher had 27 years of experience post-ERI and was replaced by a teacher with 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 to 0.13 standard deviations worse in math and 0.03 to 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 (2013) 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 5.4), 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.²²

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. 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 et al. 2002), disadvantaged schools may not have had much turnover from the ERI if their teacher workforce were too young to take advantage of the program.

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.²³ 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, although it is not possible to statistically differentiate between the estimates for advantaged and disadvantaged schools. If there are such differences across school types, it could be driven by the fact that teachers tend to move to higher-performing and wealthier districts over the course of their careers and that this type of mobility is

²² It also is 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 (2013), 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.

²³ 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.

easier for better teachers. Alternatively, Feng et al. (2012) shows that the return to experience in Florida and North Carolina is lower in high-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 in different schools. For example, if low-income schools only were able to hire novice teachers to replace retirees but high-income schools hired mid-career teachers, it would reconcile our results with the previous literature on teacher experience and teacher quality. In an effort to disentangle what might be driving any differences, in Table 8, we present estimates by school type of the effect of the ERI 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.²⁴ In order to shed some light into 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 only are 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 slightly, by between -0.09 and -0.22 for each teacher with 15 or more years of experience pre-ERI, but this represents a small decline in relative terms (just 2 percent).²⁵ However, this decrease is commensurate with a small

²⁴ In particular, these results differ from those of the 1996 California class size reduction policy, which had the effect of switching many experienced teachers from low-SES to high-SES schools (Jepsen and Rivkin, 2009). The differences in teacher sorting in response to these two programs is an interesting topic for future research.

²⁵ The effect of ERI on the total number of teachers is found by summing the coefficients in Table A-1 in each row.

decrease in enrollment at these low-resource schools; student-teacher ratios barely change as a function of the number of pre-ERI experienced teachers in the lower-performing and lower-resource schools, as shown in Table 8. Larger test score effects in the more disadvantaged schools are likely driven by some combination of fewer productive teachers retiring and their replacements having more previous experience (rather than being novices). Our data and empirical design do not let us distinguish between these two theories, however.

5.4. *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 ERI programs, or teacher retirement more generally, might lead to increases in student achievement. Though our baseline estimates 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 change negligibly 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.²⁶

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

²⁶ 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.

experience pre-ERI is negative but is not statistically significantly different from zero. These 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, across all schools, there was a 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 of early-mid career teachers switching into tested grades in the top quartile of schools, the increase was larger in the lowest quartile of schools, 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 nonnegative 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 directly compare estimates with the baseline ones, but we can use information in the data to do so indirectly. The coefficients in the second row of Table 10 correspond to an increase from no experienced teachers to all experienced teachers (i.e. from 0 to 1) in the pre-treatment period. On average in the data, an increase in one experienced teacher is related to a 0.11 increase in the proportion of experienced teachers. For math and English teachers specifically, an increase of one experienced teacher is related to an increase in the proportion of experienced teachers of 0.38 and 0.31, respectively. Multiplying the estimates by these percentages produces results that are very similar to baseline for the models using all teachers. The results using this specification also 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 next 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. Both sets of results are virtually identical to the baseline results, which suggests our imputation of the missing grade information is not driving our conclusions.

In the following robustness check, we measure all pre-treatment variables using 1993 data rather than data from all pre-treatment years. The results are very similar to baseline. We also control for the number of teachers in 1993 with 1-5 years of experience (separately), interacted with the *Post* indicator. This specification controls for any positive shock to test scores in schools with many inexperienced teachers pre-ERI that is driven by these teachers gaining experience when the program comes into place. The estimates are very similar to those in Table 4 but are somewhat larger using all math teachers and somewhat smaller using subject-specific reading teachers. Next, we control for school-by-post fixed effects that allow for unobserved school-level shocks post-ERI. This model is identified only off of those schools in which there are multiple grades per school, and thus the sample is particularly weighted to the 8th grade and rural schools. The results are similar to baseline, and are especially close in magnitude to the 8th grade results in Table 5. There is no evidence that school-specific shocks are biasing our results.

Finally, we estimate models at the district-grade rather than the school-grade level. This robustness check is motivated by the concern that some teachers may switch schools to take advantage of ERI-induced vacancies. If higher-quality teachers move from schools with lower-experience levels to those with higher-experience levels due to the ERI program, our estimates

could be biased: achievement in more heavily treated schools would increase because of teachers switching from less heavily treated schools. In this case, aggregate achievement would not necessarily increase, as schools with fewer experienced teachers would become worse due to the ERI. In order to test for these general equilibrium effects, we estimate our difference-in-difference model at the district level, under the assumption that teachers are more likely to switch schools within a district. These results are shown in the last row of Table 5 and are qualitatively very similar to our baseline estimates. While somewhat smaller in magnitude, as would be expected from the increased measurement error in the treatment intensity that comes with aggregation, they still are positive.

We also estimated equation (1) using the proportion of teachers in each district switching schools (both within and across districts) in each year as the dependent variable in order to examine such switching behavior directly. The coefficient on *Teachers* ≥ 15 is -0.0003 (with a standard error of 0.0005), off of a baseline switching proportion of 0.06. Thus, there is no evidence the ERI leads to increasing rates of school switching among teachers that could bias our estimates.

6 Conclusion

This paper presents the first evidence in the literature on the effects of teacher retirement in response to ERIs 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. 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. 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. Given available data, it is not possible to examine the underlying productivity of teachers taking up the

ERI. 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 by saving them money on teacher salaries without having a deleterious impact of retirement on student achievement. Along these lines, it is important consider the costs of this program to both school districts and the state, as well as the value of any increases in test scores that occurred because of the program. We calculate that the median teacher retires 5 years earlier than expected because of the ERI.²⁷ The median teacher therefore retires at age 55 with 27 years of service rather than waiting until age 60 with 32 years of service. Since replacing a 27-year veteran teacher with a novice one saves \$20,772 per year on average, this would result in savings per employee to the district of \$95,306 (present discounted value with a 3 percent real interest rate).²⁸ However, for each year of creditable service purchased for the ERI, the district had to pay 12% of the teacher's salary in a lump-sum payment. Since the median teacher bought 5 years of service, the median lump-sum payment made by the district was 60% of a teacher's salary, or \$26,493. Thus, the median net PDV to the district per teacher of the ERI was \$68,812. Summed across the approximately 8,000 teachers who took up the program, the ERI program resulted in total savings to IPS school districts of \$550.5 million. These cost savings were likely at the forefront of the policy discussion surrounding the ERI.

However, the teachers who retire early because of the ERI will receive pension benefits for more years than if they had retired normally. In Illinois, it is the state, rather than the districts, that must make up the increased costs to the pension system of these ERI retirees. The median ERI-retiree receives \$115,677 of increased benefits, in present value terms, from retiring five years earlier. The teacher must also pay a lump-sum amount equal to 4% of her salary per year purchased. Accounting for the district and teacher lump-sum payments, the net cost to the pension system was \$80,352 per ERI retiree, or \$642.8 million to the pension fund. Therefore, the total cost to the state's taxpayers – the sum of the net benefit to the districts and the net cost to the pension fund – is approximately \$92.3 million (PDV). Since there were approximately 1.8 million students in IPS in 1993, this represents a cost per student of \$51.

A core conclusion that stems from these costs calculations is that even when an ERI program creates substantial savings to school districts by reducing teacher wage bills, it still can cost the state money through higher pension payments. This was the case in Illinois, and it is plausible that this occurred because lawmakers did not fully understand this tradeoff. In addition to this likely unintended cost, the central focus of this paper is in showing that the ERI had a possibly unintended benefit in raising test scores. It therefore is worth considering the cost-

²⁷ For more information on the parameters in this cost-benefit analysis, see Appendix B.

²⁸ Because wages rise with tenure, we use the actual salaries of 1 to 5 year and 27 to 31 year teachers and calculate the 5-year present discounted value of the differences in salaries in each of the 5 years.

effectiveness of the policy as a tool for improving achievement, even though this may not have been its initial intention.

Our confidence intervals in Table 4 suggest that the average effect of the ERI on test scores was between -0.002 and 0.029, with an average effect across specifications of 0.010. Using the average point estimate and cost estimate, then, suggests that taxpayers paid \$51 per student in return for a one percent of a standard deviation increase in test scores. At worst, the taxpayers of Illinois paid \$51 per student to decrease test scores by 0.002 of a standard deviation. To put this cost-benefit ratio in perspective, a similar-sized improvement in student test scores resulting from class size reductions would cost about \$96 more per student, on average, than the ERI.²⁹ While conducting cost-benefit analyses and comparisons across interventions is inherently difficult, the comparison is informative of the potential relative efficiency of ERIs as interventions to increase test scores. Such policies may provide a potential means to save districts money without hurting student achievement. And, with properly set prices, it is possible for such a program to save tax-payers money as well.

²⁹ As reported in Kreuger (2003), the Tennessee STAR class size reduction cost \$7,660 per student for a 0.2 standard deviation increase in student-level test scores. As discussed above, school-grade test scores can be translated to student-level test scores by dividing by a factor of 2.6. Assuming a linear relationship between the spending on and returns to class size reductions, the class size reduction would require spending \$147 per student (\$7,660 divided by 52) to obtain the same sized increase in test scores we see with the ERI.

REFERENCES

- Aurriemma, 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>.
- 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. 2013. "The Link between Pensions and Retirement Behavior: Lessons from California Teachers." *Journal of Public Economics*. 98:1-14.
- Chingos, Matt and Martin West. 2012. "Do More Effective Teachers Earn More Outside of the Classroom?" *Education Finance and Policy* 7(1): 8-43 (2012).
- Clotfelter, C.T., Ladd, H.F., & Vigdor, J.L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, 41(4), 778-820
- 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. 2012. "Comparison of the Value Added of Teachers in High-Poverty Schools and Teachers in Lower-Poverty Schools." *Journal of Urban Economics*. 74: 104-122.
- Fitzpatrick, Maria D. 2013. "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.
- Goldhaber, Dan, and Michael Hansen. 2010. "Using Performance on the Job to Inform Teacher Tenure Decisions." *American Economic Review: Papers and Proceedings*, 100(2): 250-55.

- 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.
- Jepsen, Christopher and Steven Rivkin. 2009. "Class Size Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size." *Journal of Human Resources* 44(1): 223-250.
- 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.
- Krueger, Alan. 2003. "Economic Considerations and Class Size" *The Economic Journal*. 113(485): F34-F63.
- 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.
- Papke, Leslie E. 2005. "The Effects of Spending on Test Pass Rates: Evidence from Michigan." *Journal of Public Economics* 89(5-6): 821-839.
- Papay, John and Matt Kraft. 2011. "Productivity Returns to Experience in the Teacher Labor Market: Methodological Challenges and New Evidence on Long-Term Career Growth."

Manuscript, Harvard School of Graduate Education.

http://scholar.harvard.edu/files/mkraft/files/papay__kraft_-_productivity_returns_to_experience_in_the_teacher_labor_market_-_nov_2011.pdf

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

Rockoff, Jonah E. 2004. "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. 2013. "The Dynamics of Teacher Quality." *Journal of Public Economics*. 100: 61-78.

Table 1. Descriptive Statistics

Variable	Full Sample		3 rd Grade		6 th Grade		8 th Grade	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Number With 15 or More Years of Experience, Pre-ERI	5.19	5.24	2.27	1.48	5.07	4.57	9.76	6.40
Total Number of Teachers in Grade, Pre-ERI	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 Years of Experience, Pre-ERI	1.48	1.16	1.59	1.02	1.47	1.31	1.33	1.13
Total Number of Math Teachers in Grade, Pre-ERI	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 Years of Experience, Pre-ERI	1.80	1.36	1.60	1.02	1.78	1.42	2.14	1.64
Total Number of English Teachers in Grade, Pre-ERI	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. Exit Likelihoods by Experience Level and the Experience Distributions of Exiters, Before, During and After the ERI Program

Years of Experience	<u>Pr(Leave)</u>			<u>Difference</u>		<u># of Teachers</u>
	1990-1993 (1)	1994-1995 (2)	1996-1997 (3)	(2)-(1) (4)	(3)-(1) (5)	(6)
≤ 1	0.091	0.066	0.103	-0.025	0.012	11482
2	0.066	0.060	0.077	-0.006	0.011	11472
3	0.060	0.057	0.053	-0.003	-0.007	10995
4	0.053	0.045	0.063	-0.008	0.010	9900
5	0.056	0.060	0.059	0.004	0.003	9125
6-9	0.042	0.045	0.049	0.003	0.007	31609
10-14	0.031	0.040	0.032	0.009	0.001	36944
15-19	0.024	0.044	0.024	0.020	0.000	42754
20-24	0.032	0.062	0.019	0.030	-0.013	44181
25-29	0.051	0.129	0.023	0.078	-0.028	29694
30-34	0.111	0.367	0.054	0.256	-0.057	11951
35-39	0.253	0.407	0.169	0.154	-0.084	2835
≥ 40	0.281	0.362	0.238	0.081	-0.043	521

Years of Experience	<u>Experience Distribution of Exit</u>			<u>Difference</u>		<u># of Teachers</u>
	1990-1993 (1)	1994-1995 (2)	1996-1997 (3)	(2)-(1) (4)	(3)-(1) (5)	(6)
≤ 1	0.080	0.035	0.143	-0.045	0.063	11482
2	0.055	0.030	0.112	-0.025	0.057	11472
3	0.045	0.030	0.062	-0.015	0.017	10995
4	0.038	0.022	0.063	-0.016	0.025	9900
5	0.038	0.026	0.057	-0.012	0.019	9125
6-9	0.105	0.067	0.153	-0.038	0.048	31609
10-14	0.104	0.065	0.096	-0.039	-0.008	36944
15-19	0.104	0.086	0.076	-0.018	-0.028	42754
20-24	0.128	0.138	0.072	0.010	-0.056	44181
25-29	0.112	0.201	0.069	0.089	-0.043	29694
30-34	0.117	0.225	0.048	0.108	-0.069	11951
35-39	0.062	0.067	0.037	0.005	-0.025	2835
≥ 40	0.012	0.010	0.012	-0.002	0.000	521

Source: Illinois Teacher Service Record and school-level Illinois State Board of Education data, 1990-1997.

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 Pr(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. The second panel shows the distribution of experience among those exiting in the previous school year. Each of the first three columns sums to one in these estimates. The # of Teachers shows the total number of teacher observations in each experience group throughout the analysis sample.

Table 3. OLS Estimates of the Effect of the Early Retirement Incentive Program on Teacher Composition and Pupil-Teacher Ratios

Dependent Variable	Full Sample	3 rd Grade	6 th Grade	8 th Grade
Number of Teachers with 15 or More Years of Experience Exiting in the Previous School Year	0.078** (0.013)	0.081** (0.011)	0.085** (0.024)	0.075** (0.019)
Dependent Variable Mean	0.239	0.108	0.234	0.444
Average Teacher Experience in Grade	-0.405** (0.040)	-1.088** (0.103)	-0.472** (0.060)	-0.241** (0.049)
Dependent Variable Mean	15.52	15.39	15.25	16.09
Number of New Teachers	0.073** (0.014)	0.076** (0.019)	0.088** (0.017)	0.063** (0.022)
Dependent Variable Mean	0.368	0.156	0.405	0.635
Pupil-Teacher Ratio	-0.009 (0.035)	0.183 (0.196)	-0.002 (0.060)	0.020 (0.027)
Dependent Variable Mean	14.95	19.46	13.36	10.35

Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1995 combined with school-level Illinois State Board of Education data on test scores.

Notes: The table shows estimates of β_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-by-grade fixed effects, and grade-by-year fixed effects. All estimates except for pupil-teacher ratios also contain a quadratic in school-grade-year enrollment. 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

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

Dependent Variable Mean	0.117	0.048	0.105	0.238
40+	-0.002 (0.002)	0.0002 (0.003)	-0.0003 (0.0004)	-0.003 (0.002)
Dependent Variable Mean	0.017	0.011	0.017	0.024

Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1995 combined with school-level Illinois State Board of Education data on test scores.

Notes: The table shows estimates of β_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-by-grade fixed effects, grade-by-year fixed effects and a quadratic in school-grade-year enrollment. 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

Independent Variable	All Teachers		Subject-Specific Teachers	
	Math	Reading	Math	Reading
Post*Number \geq 15 Years of Experience, Pre-ERI	0.003 (0.004)	0.009** (0.003)	0.013* (0.008)	0.013* (0.007)
Post*Total Number of Teachers, Pre-ERI	0.0001 (0.002)	-0.004* (0.002)	-0.004 (0.006)	-0.005 (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, 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-by-grade fixed effects, grade-by-year fixed effects and a quadratic in school-grade-year enrollment. 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

Independent Variable	All Teachers		Subject-Specific Teachers	
	Math	Reading	Math	Reading
Post*Number \geq 15 Years of Experience, Pre-ERI *I(Grade=3)	0.002 (0.010)	-0.009 (0.008)	0.010 (0.013)	-0.003 (0.010)
Post*Number \geq 15 Years of Experience, Pre-ERI *I(Grade=6)	-0.00002 (0.005)	0.006 (0.004)	0.004 (0.010)	0.016* (0.009)
Post*Number \geq 15 Years of Experience, Pre-ERI *I(Grade=8)	0.005 (0.005)	0.013** (0.005)	0.032* (0.018)	0.030* (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-by-grade fixed effects, grade-by-year fixed effects and a quadratic in school-grade-year enrollment. 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

Ind. Var.: Post*Number ≥ 15		
School Type	Math	Reading
Lowest Income Schools	-0.001 (0.008)	0.014 (0.008)
All Other Schools	0.002 (0.004)	0.009** (0.004)
Lowest % White Schools	0.004 (0.007)	0.018** (0.007)
All Other Schools	0.004 (0.004)	0.006 (0.004)
Lowest Baseline Score Schools	0.010 (0.008)	0.019** (0.008)
All Other Schools	0.003 (0.004)	0.008** (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-by-grade fixed effects, grade-by-year fixed effects and a quadratic in school-grade-year enrollment. 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 Pupil-Teacher Ratios, 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 (0.034)	0.085** (0.016)	0.049* (0.026)	0.091** (0.016)	0.038 (0.029)	0.087** (0.016)	0.046* (0.028)	0.086** (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** (0.083)	-0.465** (0.037)	-0.203** (0.081)	-0.496** (0.040)	-0.214** (0.098)	-0.468** (0.037)	-0.218** (0.096)	-0.473** (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** (0.025)	0.104** (0.013)	-0.024 (0.027)	0.107** (0.014)	-0.062** (0.027)	0.107** (0.013)	-0.008 (0.005)	0.108** (0.013)
Dep. Var. Mean	0.248	0.395	0.342	0.377	0.312	0.384	0.314	0.384
Pupil-Teacher Ratio	-0.018 (0.150)	-0.011 (0.035)	0.044 (0.087)	0.006 (0.040)	-0.001 (0.113)	-0.002 (0.037)	0.018 (0.109)	-0.004 (0.037)
Dep. Var. Mean	20.61	13.99	18.51	14.14	19.30	14.00	19.26	13.97

Source: Authors' estimation of equation (1) using Illinois Teacher Service Record data from 1990 through 1995 combined with school-level Illinois State Board of Education data on test scores.

Notes: The table shows estimates of β_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-by-grade fixed effects, grade-by-year fixed effects and a quadratic in grade-school-year enrollment. The pupil-teacher ratio estimates do not control for enrollment. 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

Independent Variable	All Teachers – Excluding 8 th Grade		All Teachers – Including Higher Grade Teachers	
	Math	Reading	Math	Reading
Post*Number \geq 15 Years of Experience, Pre-ERI	0.001 (0.005)	0.002 (0.004)	0.001 (0.005)	0.004 (0.004)
Post*Total Teachers, Pre-ERI	0.001 (0.003)	0.0004 (0.003)	0.002 (0.003)	-0.00002 (0.003)
Post*Number Other \geq 15 Years of Experience, Pre-ERI			-0.005 (0.005)	-0.006 (0.004)
Post*Total Other Teachers, Pre-ERI			-0.001 (0.003)	0.002 (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-by-grade fixed effects, grade-by-year fixed effects and a quadratic in school-grade-year enrollment. "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.

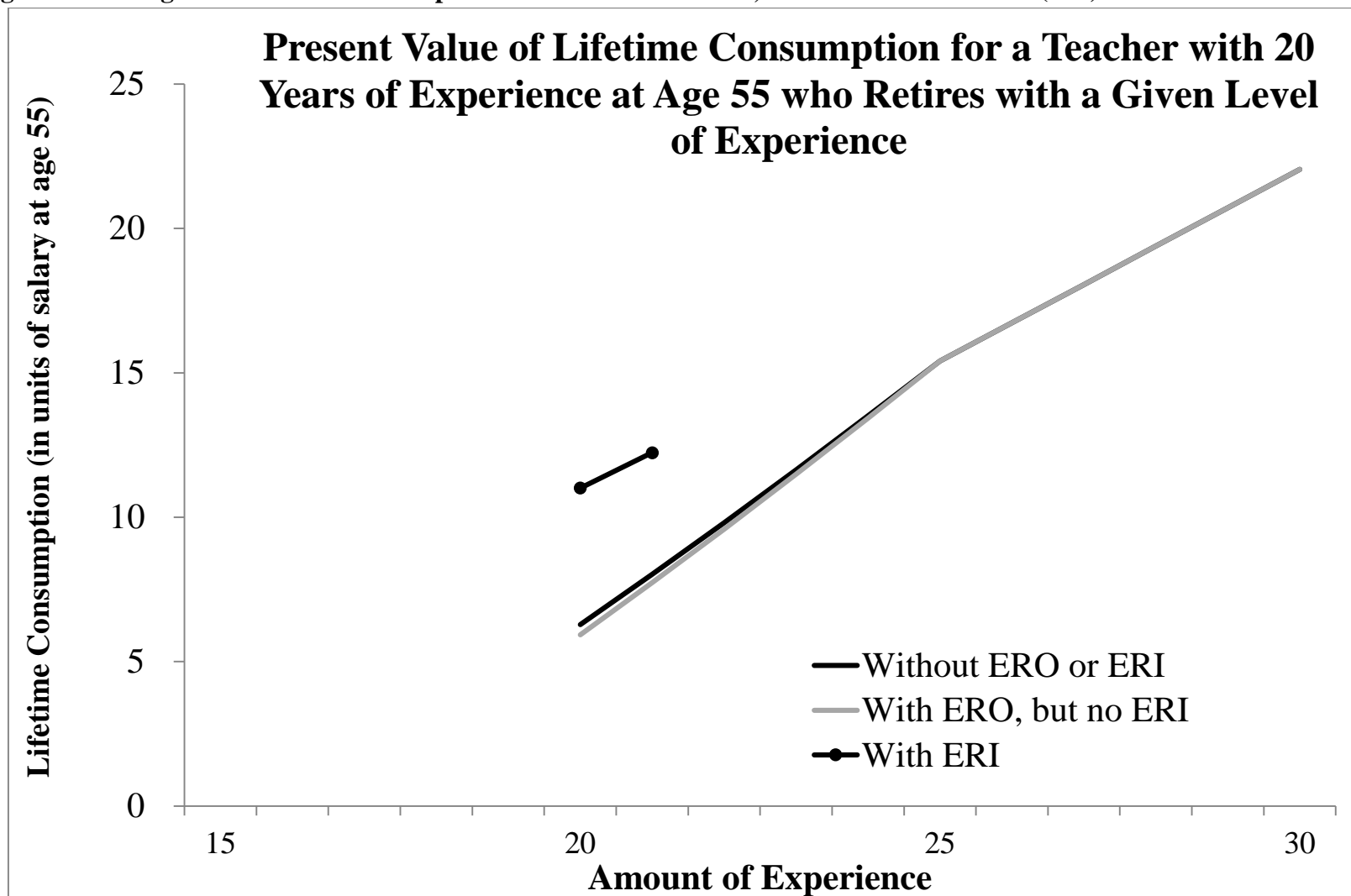
Table 10. Robustness Checks

Specification	All Teachers		Subject-Specific Teachers	
	Math	Reading	Math	Reading
Post* Number \geq 20 Years of Experience, Pre-ERI	0.002 (0.003)	0.009** (0.003)	0.008 (0.008)	0.008 (0.007)
Post* Percentage \geq 15 Years of Experience, Pre-ERI	0.030 (0.039)	0.063* (0.035)	0.007 (0.019)	0.007 (0.019)
Post* Number \geq 15 Years of Experience, Pre-ERI (Excluding Chicago)	0.004 (0.004)	0.009** (0.003)	0.015* (0.007)	0.013* (0.007)
Post* Number \geq 15 Years of Experience, Pre-ERI (Excluding Missing)	0.002 (0.004)	0.009** (0.003)	0.013* (0.008)	0.013* (0.007)
Using 1993 as Pre-Treatment Year	0.006* (0.003)	0.009** (0.003)	0.012* (0.007)	0.010 (0.006)
Controlling for # Inexperienced in 1993 * Post	0.008** (0.004)	0.008** (0.004)	0.018** (0.008)	0.006 (0.006)
School-Post Fixed Effect	0.007 (0.007)	0.003 (0.005)	0.032** (0.013)	0.030** (0.013)
District-level Estimates	0.002 (0.003)	0.002 (0.004)	0.008 (0.007)	0.004 (0.006)

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.

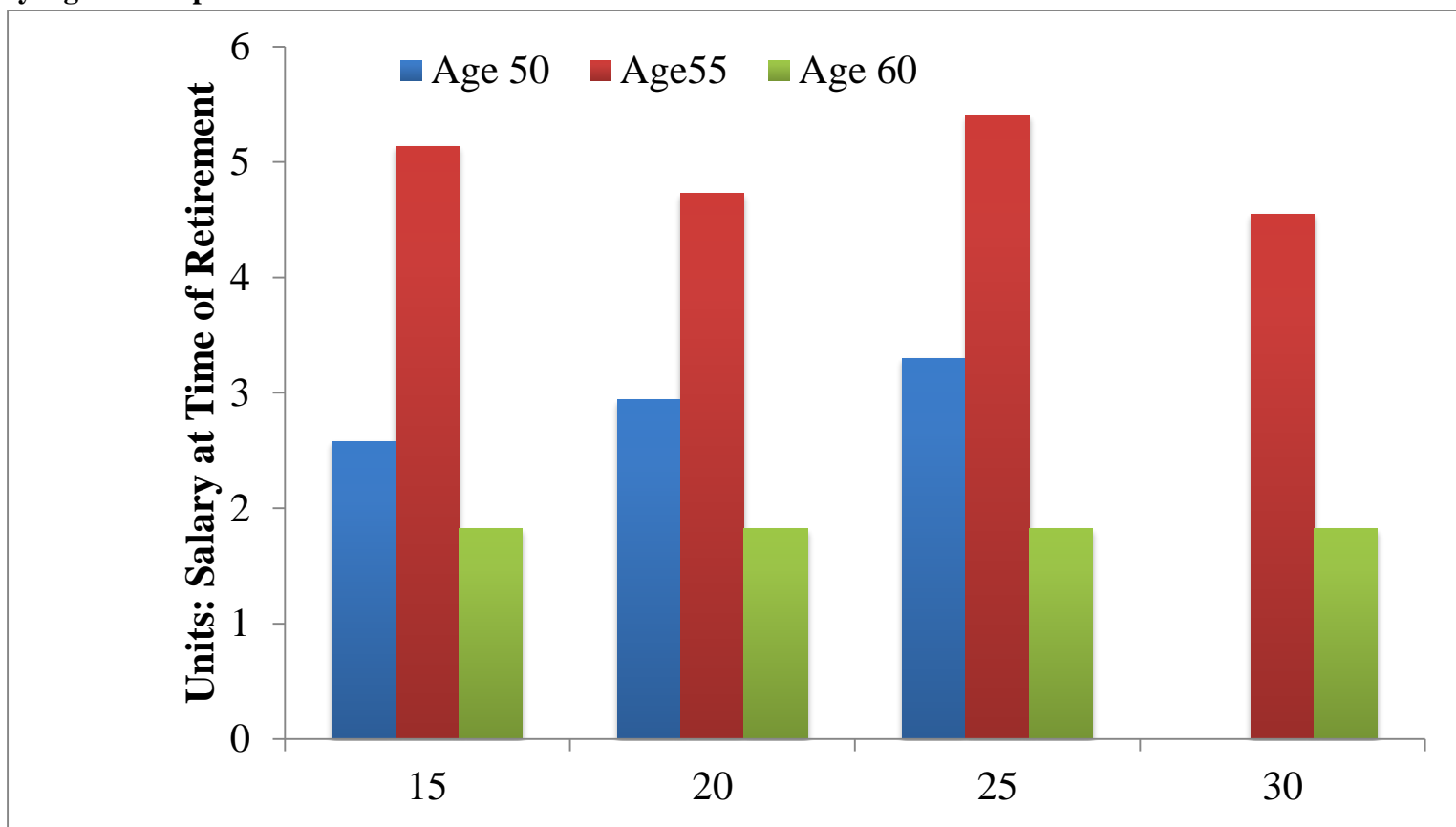
Notes: Each column presents results from a separate regression. Unless otherwise noted, estimates show the coefficient on the *Post*Number \geq 15* variable. Teachers who teach multiple grades are included in the estimates for each grade in which they teach. Teachers who teach in self-contained classrooms are assumed to teach both math and English. All estimates aside from those in the last row include controls for the demographic variables included in Table 1, school-by-grade fixed effects, grade-by-year fixed effects and a quadratic in school-grade-year enrollment. The estimates in the last row contain demographic variables, district-grade fixed effects, grade-year fixed effects and a quadratic in district-year enrollment. The estimates in all but the second row 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 (clustered at district-grade in the final row): ** indicates statistical significance at the 5% level and * indicates statistical significance at the 10% level.

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



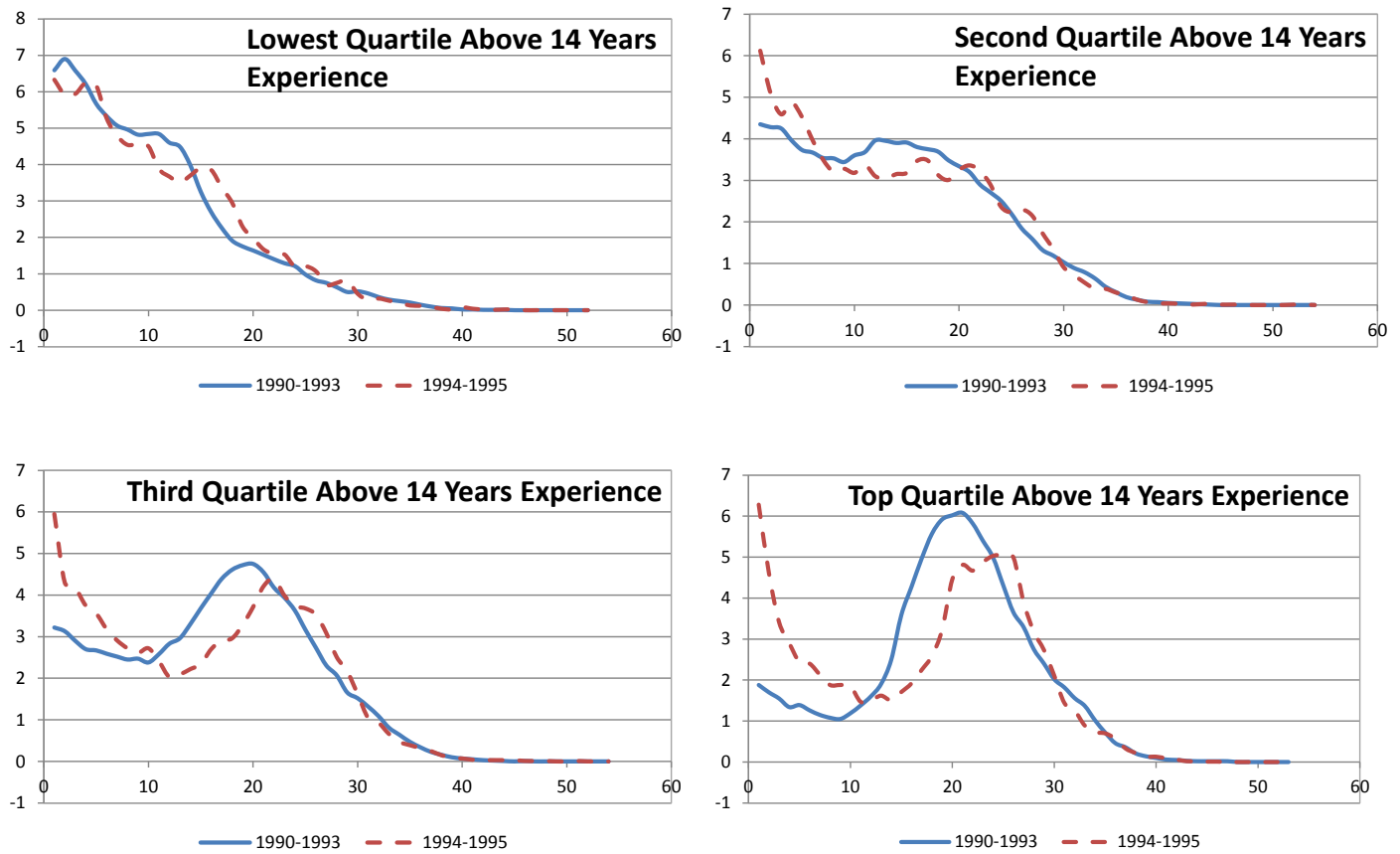
Notes: Based on authors' calculations. The vertical axis measures the present value of consumption for a teacher who is age 55 and has 20 years of experience when the ERI becomes available. The vertical axis presents the present discounted value of lifetime consumption associated with retirement with 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 nominally and we assume a real interest rate of 3 percent. The present value of consumption is measured at age 55 and is presented in multiples of salary at age 55.

Figure 1.B. Increase in Lifetime Consumption Relative to Normal Retirement of Representative Teachers in IPS Due to ERI (5+5), by Age and Experience



Notes: Based on authors' calculations. The vertical axis measures the ERI-induced increase in the present value of consumption for a teacher who is of the age and experience level indicated when the ERI becomes available, relative to not taking up the ERI. 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 nominally and we assume a real interest rate of 3 percent. The present value of consumption is measured at age 55 and is presented in multiples of salary at age 55. No 50 year old teachers with 30 years of experience exist in the data, so we have omitted this category.

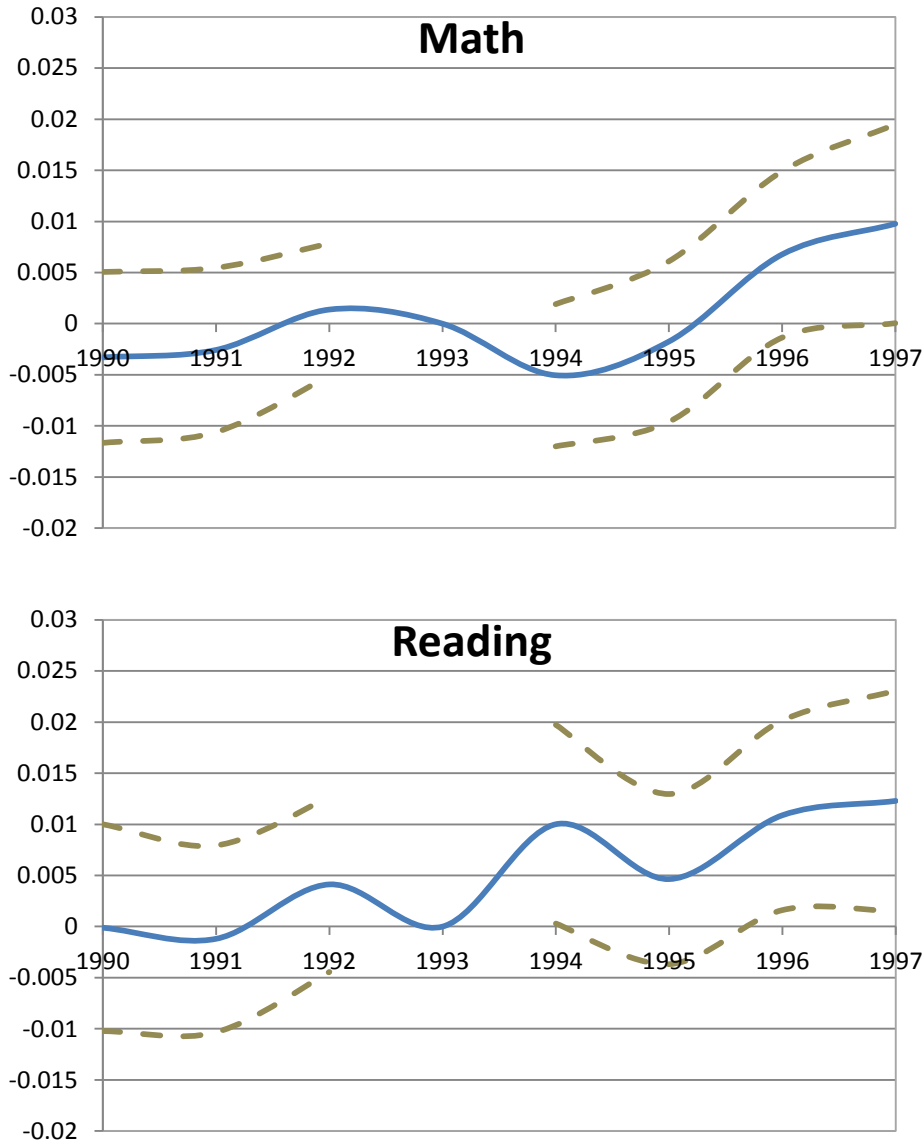
Figure 2. Distribution of Teacher Experience Pre- and Post-ERI Program, by Quartiles of the Percent of Grade-Specific Teachers with 15 or More Years of Experience Prior to 1994



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

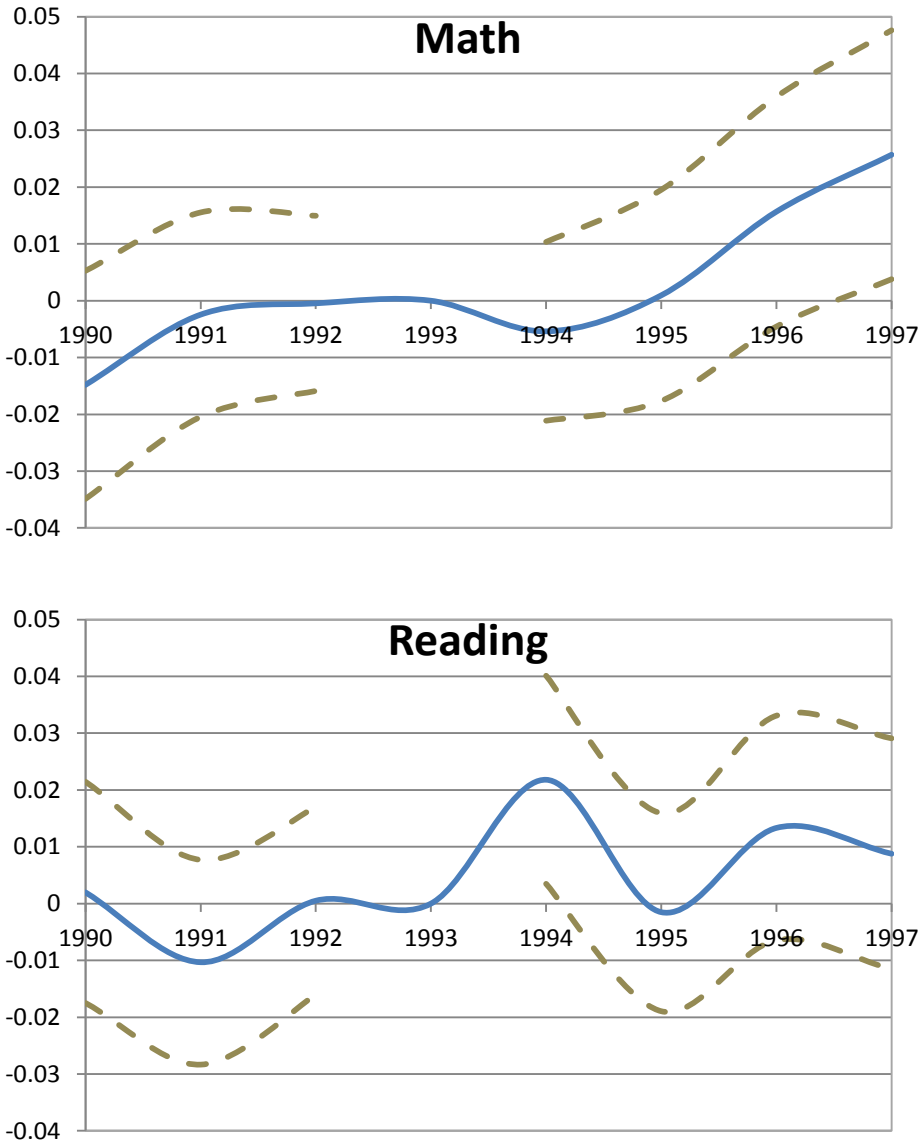
Notes: All years refer to the calendar year in which a school year ends, and distributions are based on all teachers in 3rd, 6th and 8th grade. Each panel of the figure shows experience distributions for the pre-treatment period (1990-1993) and the period immediately following the early retirement incentive program implementation (1994-1995) by the quartile of the percent of teachers with 15 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



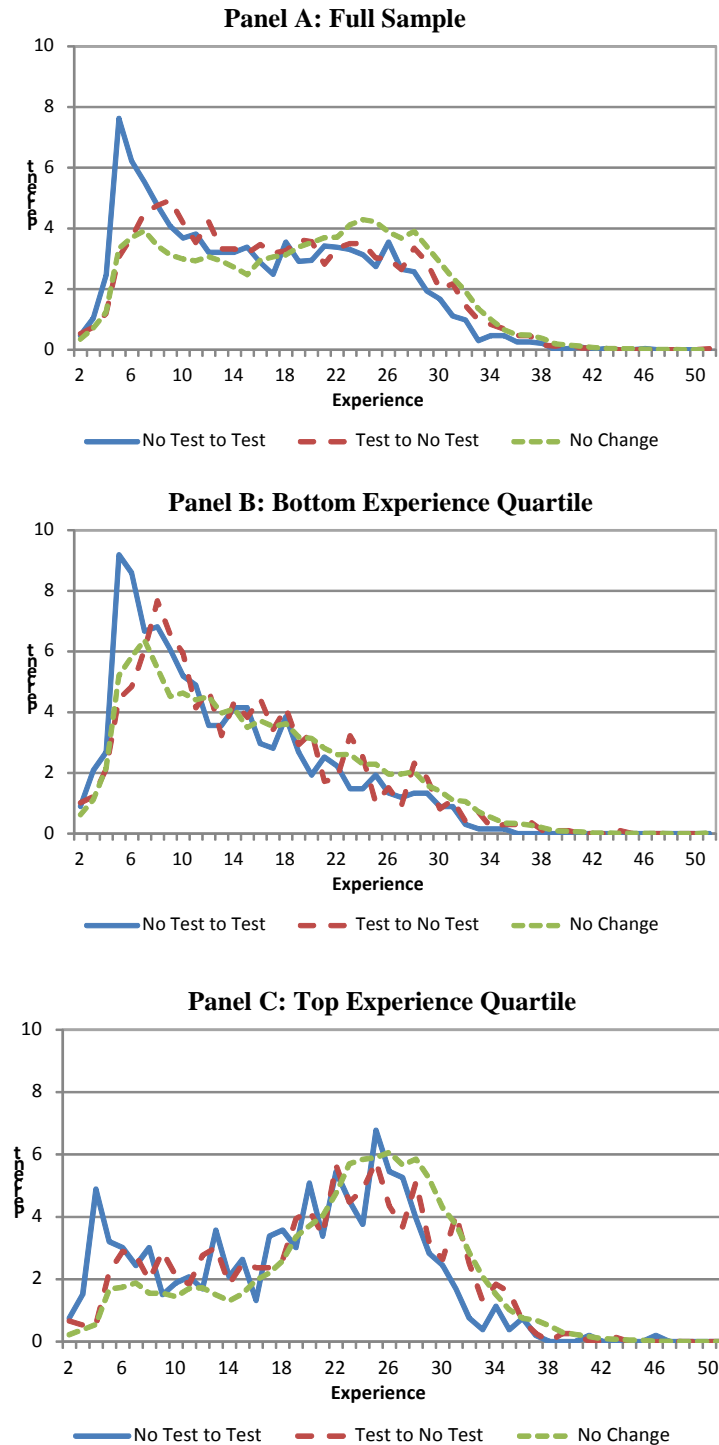
Notes: Estimates and 95% confidence intervals from interactions of year fixed effects and the number of teachers with 15 or more years of experience. Years are indexed by the calendar year in which a school year ends. The 1993 coefficient is set to zero, so there is no standard error bound for this year.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

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-by-grade fixed effects, the demographic characteristics shown in Table 1, interactions between year fixed effects and the number of subject-specific teachers in each grade pre-ERI, and a quadratic in school-grade-year enrollment

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.

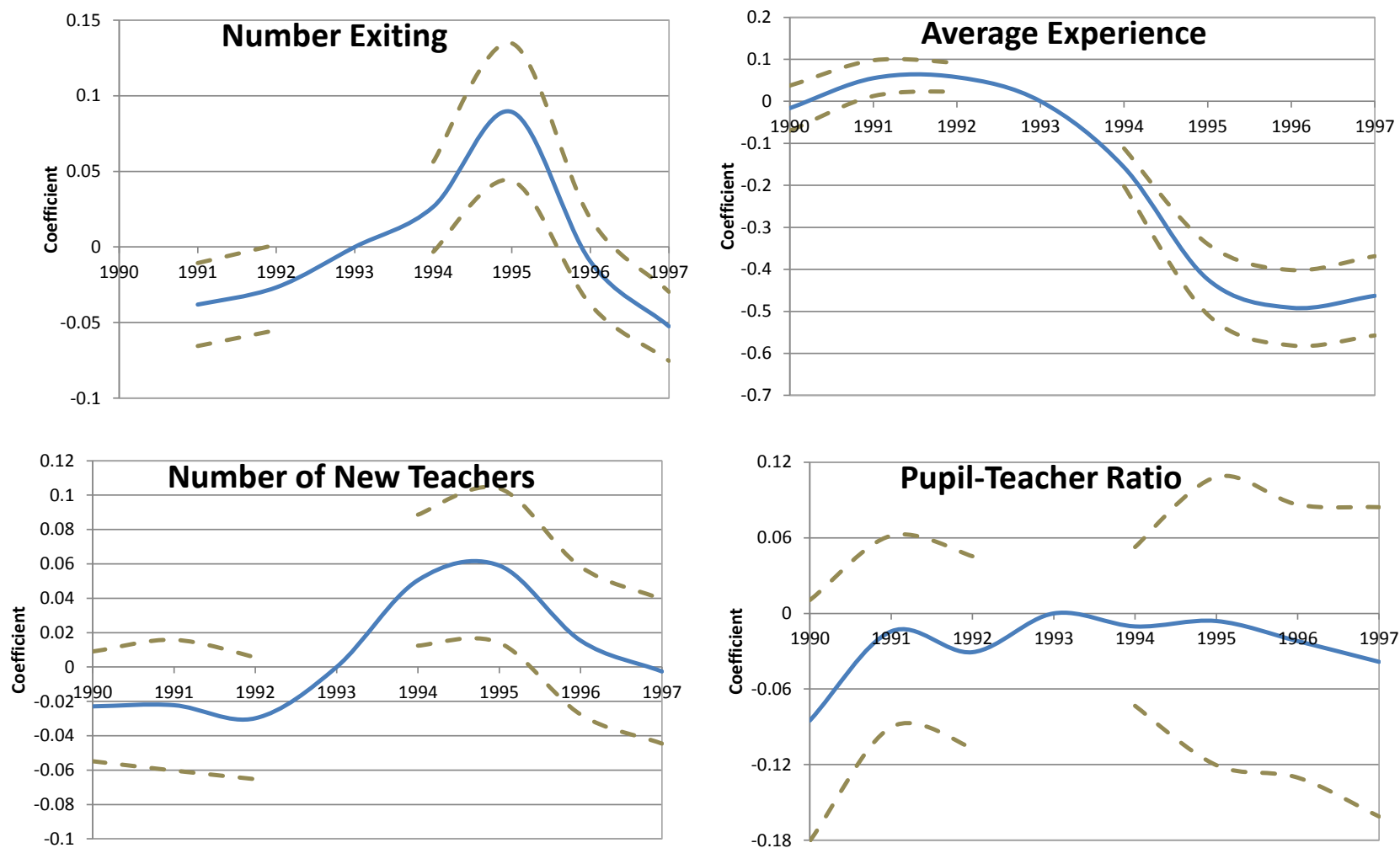
Appendix A: Supplementary Tables and Figures

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

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

Notes: Data come from the 1990-1995 TSR combined with school-level Illinois State Board of Education data on test scores. The table shows estimates of β_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-by-grade fixed effects, grade-by-year fixed effects and a quadratic in school-grade-year enrollment. 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-by-grade fixed effects, the demographic characteristics shown in Table 1, and interactions between year fixed effects and the number of teachers in each grade. The pupil-teacher ratio estimates do not control for enrollment.

Appendix B. Cost-Benefit Calculations

The cost-benefit calculations reported in Section 6 depend on parameters we have chosen using our data. These are described in turn:

How Much Earlier Do Teachers Retire Because of the ERI?: We first calculate the number of teachers with each level of experience in the pre-ERI period and the exit rates of teachers with each given level of experience in both the pre-ERI and during-ERI periods. Using the number of teachers with each level of accrued experience in the pre-ERI period as the baseline, for each given level of starting experience (e.g. 15 years of service, 16 years of service, etc.), we create conditional density functions for the fraction of teachers that exit the system with each passing year. We perform this calculation for both the period pre-ERI and the period over which the ERI was offered. We then examine the median of this distribution before and during the ERI for each level of experience of at least 15 years. By this measure, the median teacher with 15 years of service or more is retiring with 32 years of experience in the pre-ERI period and retires 5 years earlier, with 27 years of experience in the ERI period.

Median Retirement Age Pre-ERI: Using the pre-ERI level of experience for the median exiting teacher described above, we estimate the median retirement age assuming a teacher starts her career at age 28 and does not have any employment breaks due to childrearing or other reasons. The median experience level of a retiring teacher with 15 or more years of experience is 32 years, which would make her 60 years old under this assumption. We further assume that pensions are paid out until age 87, on average. This means teachers receive their pension payments for 27 years pre-ERI and for 32 years post-ERI.

Salary Levels: We use average teacher salaries of all teachers with a given level of experience in 1992, the year before the ERI was introduced.

Lump-Sum Payment Calculations: Lump-sum payments are made by the district and teacher, separately, for each teacher taking up the ERI. Districts pay 12% of the highest salary, which we assume is the salary in 1992, for each year purchased. Since the median teacher purchases 5 years, we assume districts pay 60% of the 1992 salary of a teacher with 27 years of experience. Teachers must pay 4% of this salary for each year purchased, which is 20% of the 1992 salary for a teacher with 27 years of experience if she purchases 5 years of experience.

Interest Rate: We assume an interest rate throughout of 3%.

Percentage of Salary Paid: The pension system includes a formula for the cumulative percentage of salary paid by the pension: 1.67% of salary for the first 10 years, 1.91% for

the next 10 years, 2.1% for the following 10 years, and 2.3% thereafter. A teacher with 32 years of experience, which is the median purchased experience level of teachers retiring under the ERI, will receive 61.4% of her highest salary in benefits every year ($=10 \times .0167 + 10 \times .0191 + 10 \times .021 + 2 \times .023$). We use the average salary of a teacher with 27 years of experience in 1992 to calculate these benefits. Note that the median teacher retiring under ERI has 27 years of experience but is treated as if she has 32 years due to the ERI program. This is why the experience levels used to calculate the payout rate and the base salary differ.