Implications of Teacher Tenure on Teacher Quality and Student Performance in North Carolina

Dana E. Fenster*

Professor Hugh Macartney, Faculty Advisor

Duke University
Durham, North Carolina
2014

* The author can be contacted at danafenster@gmail.com. Ms. Fenster is currently a Teach for America corps member and will be beginning her first year as a high school mathematics teacher in Detroit this fall.

Acknowledgements

I would like to thank Professor Hugh Macartney, my faculty advisor, without whom this project would not have been possible. I also thank Professor Kent Kimbrough for his continued guidance and support throughout this project. Remote access to the data for this research was generously provided by the North Carolina Education Research Data Center (NCERDC).

Abstract

This paper examines the relationship between teacher tenure and teacher quality in North Carolina, measured via student performance on the state End of Grade (EOG) standardized tests. After presenting a comprehensive synopsis of the current teacher tenure policy, I use data from the North Carolina Education Research Data Center (NCERDC) to compare demonstrated teacher effectiveness across the tenure bubble, defined as one to eight years of teaching experience within the same district. Ultimately, I find that there is a significant jump in average teacher quality at the tenure cutoff, suggesting that tenure policy is effective in retaining high quality teachers while removing those who are ineffective.

JEL classifications: I21; J24; J41; M5

Keywords: Labor Economics; Economics of Education; Teacher Tenure

I. Introduction

It is widely accepted by educators, policymakers, and researchers that teacher quality plays a significant role in determining students' academic success and long-term economic prosperity. A highly effective teacher improves both students' academic learning in the short-term and their long-term quality of life. Using the teacher value-added measure, one proxy for teacher quality, Chetty, Friedman, and Rockoff (2013) demonstrate that students with a highly effective teacher not only improve their end-of-year test scores in that year, but are also more likely to graduate high school, matriculate to college, and earn higher incomes (as compared to their peers with a low value-added teacher). A single year with a high value-added teacher is shown to increase a student's cumulative lifetime income by an average of $80,000, or nearly 2% of total lifetime earnings.[1] These effects are cumulative, with each additional year of high quality instruction continually improving student learning and long-term success.[2]

Given the demonstrated importance of quality teachers, much attention has been paid to the attraction and retention of highly effective teachers. One such policy, teacher tenure, is the subject of recent education reform debates. While teacher tenure provides effective teachers with job security and protection from wrongful termination, it also makes it difficult to remove teachers who are ineffective in the classroom.[3] From a labor economics lens, the job security provided by tenure policies could alter incentives for teacher behavior, thus impacting teacher effectiveness. The aim of my research is to evaluate the effects of tenure policy on teacher quality and the resulting student performance, particularly in North Carolina.

My analysis focuses solely on North Carolina for three reasons. First, the motivation for this research stems partially from the state's unique policies regarding collective bargaining and teacher tenure. Additionally, thanks to the establishment of the

---

[1] Raj Chetty, John N. Friedman, and Jonah E. Rockoff, "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood," *National Bureau of Economic Research* Working Paper 19423 (2013): 41-44.

[2] Ron Haskins and Susanna Loeb, "A Plan to Improve the Quality of Teaching in American Schools," *The Future of Children* (2007): 1-2.

[3] M.J. Stephey, "A Brief History of Tenure," *Time*, November 17, 2008, http://content.time.com/time/nation/article/0,8599,1859505,00.html.

North Carolina Education Research Data Center (NCERDC) in 2000, North Carolina's education data is some of the most thorough in the country, making it a good site for empirical research. Finally, this research is particularly relevant in North Carolina right now, given that the state government recently passed massive teacher tenure reform that will gradually eliminate tenure for the majority of teachers in the next five years (but has yet to go into effect).[4] Thus, it is natural to evaluate the existing policy before the upcoming changes take place.

Before examining this from an empirical angle, this paper presents background information on North Carolina's teacher tenure and collective bargaining laws from a policy standpoint. This discussion serves to contextualize my empirical findings within North Carolina's unique educational landscape. Then, using data from the NCERDC, I focus on the average quality of teachers across the tenure bubble, both before and after they receive tenure. Ultimately, I am interested to see if there are empirically observable differences in quality between teachers who do and do not have tenure. As will be discussed in further detail later in the paper, the goal of an optimal teacher hiring practice is to attract and retain high quality teachers while removing those who are ineffective in the classroom. I am investigating whether or not North Carolina's existing tenure policy achieves this goal.

Specifically, I examine how demonstrated teacher effectiveness, measured via student performance, differs between tenured and non-tenured teachers. To address this question, I use a measure of teacher quality, called teacher fixed effects, to compare the average teacher quality across the tenure bubble. As teachers in North Carolina are eligible to receive tenure after four years of teaching, this tenure bubble is defined as years one through eight. I am interested to see if there are any changes in demonstrated teacher quality, measured via student performance, before (years one through four) and after (years five through eight) tenure is granted. Given that student test score performance is one accepted measure of teacher effectiveness, this allows me to draw preliminary conclusions about the relationship between teacher tenure and teacher quality and effectiveness in North Carolina.

---

[4] Lynn Bonner, "Proposed NC Budget Would End Teacher Tenure, Pay Tuition Vouchers," *News Observer*, July 21, 2013, http://www.newsobserver.com/2013/07/21/3046589/state-budget-would-end-teacher.html.

This paper begins with an overview of the current landscape for teachers in North Carolina, summarizing teacher salary, tenure, and evaluation policies. It then presents a review of relevant literature on contract incentives, teacher experience, and teacher behavior. Moving into the empirical analysis, an explanation and summary of the primary variables used in the analysis follows. Finally, a non-parametric specification is used to compare average teacher quality across the tenure bubble, leading to preliminary conclusions about the effectiveness of North Carolina's teacher tenure policy in retaining high quality teachers while removing those who appear ineffective.

II. Background

As it stands, collective bargaining is currently prohibited in the state of North Carolina, thus rendering a formal teachers union illegal. As one would expect with a lack of collective bargaining power, teacher salaries in North Carolina are relatively low when compared with those in other states.[5] According to a 2013 report from the National Education Association (NEA), the starting salary for public school teachers in North Carolina is $30,778, the third lowest in the nation. Compared to the 2011-2012 school year, the national average starting teacher salary rose 1.5% in 2013, reaching $36,141, while North Carolina's wage stayed constant. Teacher salaries in North Carolina are consistently about 15% below the national average and in stark comparison to the nearly $50,000 paid to new teachers in states with a strong union presence (such as New Jersey).[6] Acknowledging this discrepancy, current North Carolina Governor Pat McCrory recently proposed an increase in teacher salaries. If passed, the proposal will raise starting teacher pay to $35,000 (still below the current national average) by the 2015-16 school year, an almost 14% percent increase from current wages.[7]

Despite relatively low salaries, under the current law teachers in the state are eligible to receive tenure after four consecutive years of teaching in the same public

---

[5] Barry T. Hirsch, David A. Macpherson, and John V. Winters, "Teacher Salaries, State Collective Bargaining Laws, and Union Coverage," prepared for American Economic Association Meetings, San Diego (January 6, 2013): 10.

[6] National Education Association, *2012-2013 Average Starting Teacher Salaries by State*, compiled from the Collective Bargaining/Member Advocacy Database (December, 2013).

[7] Lynn Bonner and Jane Stancill, "McCrory Seeks to Raise Teachers Starting Pay," *News Observer*, February 10, 2014, http://www.newsobserver.com/2014/02/10/3608266/mccrory-increase-teachers-base.html.

school district. A year of teaching is defined as at least 120 full-time workdays. Workdays are stated to include teacher development days (in which students are not present), but do not include vacation or sick days. To be eligible for tenure, these four full years must be completed consecutively within the same school district. The only exception to this rule is medical leave (sick, disability, or maternity), in which the part-time year does not count toward tenure eligibility but does not break the chain of consecutive teaching years. For example, if a teacher completes two full (120 day) school years in a district, followed by a year of less than 120 days due to vacation time, followed by two more full years (all in the same district), they are not yet eligible for tenure. However, if a teacher completes two full years in a district, followed by a year of less than 120 days due to medical leave, followed by two more full years in that same district, they are eligible for tenure. Furthermore, as tenure is awarded at the district level, years of teaching experience do not necessarily transfer with teachers who switch school districts (this decision is left to the individual districts). [8]

Using the above criteria, at the conclusion of each school year the district superintendent submits a list of all teachers that are eligible for tenure to the district board of education. For each teacher on this list, the superintendent includes his or her recommendation for whether or not the teacher should be awarded tenure. There are no common, explicitly stated criteria for this recommendation; rather they are based on teacher evaluations by the school over the teacher's career. For each teacher on this list, the board of education then votes to either award the teacher tenure or not. Teachers who are not awarded tenure are dismissed from the district.[9]

The primary benefit of tenure for teachers is not an increase in salary, but an element of job security. North Carolina classifies teachers as either probationary (years one through four) or tenured (also called career teachers). If deemed ineffective by the school or district, probationary teachers can simply be dismissed at the end of the school year by the local board of education. Tenured teachers, on the other hand, can be discharged only for specific reasons (detailed in G.S.115C-325(e)(1)) and are legally

---

[8] Robert P. Joyce, *The Law of Employment in North Carolina's Public Schools* (Chapel Hill: The University of North Carolina at Chapel Hill, 2000), 396-399.

[9] Joyce, *The Law of Employment in North Carolina's Public Schools*, 399-400.

entitled to due process if the district seeks to dismiss them.[10] Under the North Carolina Teacher Tenure Act, tenure protection is not directly tied to a teacher's specific employment, but rather to their salary. For example, a teacher may be transferred from a preferred teaching assignment to a less desirable school or subject without violation of the act, as long as their salary remains the same.[11]

Tenured teachers in North Carolina are also subject to more relaxed teacher evaluation procedures. While probationary teachers undergo frequent detailed evaluation throughout the year, tenured teachers are evaluated only during years of licensure renewal with an abridged process and rubric.[12] North Carolina's teacher evaluation strategy is cited as an assessment both of teaching and for teaching, meaning teacher evaluation is designed to help teachers continually improve their practice.[13] Thus, this policy of limited evaluation for tenured teachers has potential implications for teacher effectiveness and student performance.

As the outlined tenure and evaluation policies apply only to traditional public school teachers, my analysis does not include private schools or public charter schools. Given the nature of their funding, private institutions are not subject to these government policies and are free to create their own protocols for hiring and dismissing teachers. Though publicly funded, charter schools have similar freedoms to construct their own teacher contract and evaluation practices. Thus, it is natural to focus solely on teachers in traditional public schools.

III. Literature Review

The relationship between tenure and behavior is a subset of a broader research area in labor economics, called personnel economics, which examines how contracts and pay structure impact incentives and behavior. In a seminal paper in personnel economics, Lazear (2000) studies the relationship between performance pay and worker productivity.

---

[10] North Carolina General Assembly, *Principal and Teacher Employment Contracts*, G.S.115C-325(c-e)(1).

[11] Joyce, *The Law of Employment in North Carolina's Public Schools*, 379-382.

[12] McREL in collaboration with the North Carolina State Board of Education, *North Carolina Teacher Evaluation Process* (2009): 19.

[13] Olivia Little, Laura Goe, and Courtney Bell, "A Practical Guide to Evaluating Teacher Effectiveness," *National Comprehensive Center for Teacher Quality* (2009): 15-16.

He concludes that switching to a piece-rate pay system significantly increases worker productivity in a factory setting. More relevant to my research, Lazear finds that about half of the increase in productivity results from the firm's new ability to attract and retain the most productive workers.[14] This finding demonstrates the important role that hiring practices, such as tenure policy, play in determining overall firm productivity.

Furthermore, Lazear finds that switching to a piece-rate pay structure leads to more variation in output. This suggests that without the extra economic incentive of piece-rate pay, more productive workers refrain from working to their highest ability.[15] This observed variation in productivity is pertinent to my research, as tenured teachers may not have the same incentives to differentiate themselves from their fellow teachers as probationary teachers do. On the other hand, given that their jobs are secure, it is possible that tenured teachers are better able to experiment with new teaching methods and thus improve their practice, whereas probationary teachers feel a need to stick to the status quo in order to earn tenure. These possible complex effects of tenure policy provide further motivation to evaluate its effectiveness and implications on student performance.

Finally, it should be noted that discussion of this literature is not to suggest that a piece-rate pay structure for teachers will improve teacher productivity and effectiveness, but rather to introduce some of the previously identified ways that contracts can impact behavior.

This so-called "contract phenomenon" is not limited to a factory setting and has been applied to a variety of topics, from sports performance to business seasonality.[16] To elaborate, Heubeck and Scheuer (2003) examine how player performance varies over the course of an athlete's contract, finding that on average athletes' performance increases during a year leading up their contract renewal.[17] My research applies this phenomenon

---

[14] Edward P. Lazear, "Performance Pay and Productivity," *American Economic Review* 90, no. 5 (2000): 1346-1347.

[15] Lazear, "Performance Pay and Productivity," 1347.

[16] Paul Over, "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality," *Quarterly Journal of Economics* 113, no. 1 (1998): 149-85.

[17] Tina Heubeck and Jochen Scheuer, "Incentive Clauses in Players' Contracts in Team Sports- Theory and Practice," *German Working Papers in Law and Economics* 2003 (2003): 1-30.

to teaching, an area with a drastically different pay scale and contract structure than professional sports. However, using a similar framework, I aim to discern the relationship between teacher tenure contracts and teacher effectiveness.

In the area of education, there has been much research on the relationship between teacher experience and teacher quality, as measured in terms of student performance. While earlier literature found that beyond the first three years of teaching, the returns to teacher experience are insignificant,[18] Wiswall's (2013) recent paper concludes that there continues to be a statistically significant relationship between teacher quality and teacher experience. Thus, in the simplest terms, he shows that teacher experience beyond the third year still matters in terms of teacher effectiveness and the resulting student performance. These new results are justified by the fact that unlike previous research, Wiswall is able to account for teachers leaving the profession after their first few years of teaching. In fact, he also finds that high quality teachers are more likely to exit teaching, as they tend to have a higher general skill level and are thus able to switch to more lucrative careers.[19] This suggests that tenure policy could be important in improving not just overall teacher retention, but specifically the retention of high quality teachers, by offering a job security incentive in place of a higher salary. The policy implications of Wiswall's research are twofold. Policymakers should aim to create a policy that allows for the dismissal of low quality teachers while incentivizing those who are more effective to continue teaching. My empirical research builds off of this, serving as a preliminary test to how effective tenure policy is in removing low quality teachers while retaining those who are highly effective.

On the topic of evaluating the relationship between contract structure and teacher behavior, Jacob's (2010) paper examines the effects of a new teacher tenure policy on teacher productivity in the Chicago Public Schools (CPS). He focuses on a policy enacted by CPS in 2004, which enables principals to dismiss teachers with less than five years of experience (deemed probationary teachers by CPS) without the due process and

---

[18] Douglas N. Harris and Tim R. Sass, "Teacher training, teacher quality, and student achievement, *Journal of Public Economics* 95 (2011): 798-812.

[19] Matthew Wiswall, "The Dynamics of Teacher Quality," *Journal of Public Economics* 100 (2013): 61-78.

resulting bureaucratic difficulty that traditionally comes with dismissing tenured teachers. Using a difference-in-difference estimator to measure changes in teacher absences before and after the policy change, Jacob finds that teacher absences fell by 10-20% following the new policy's introduction.[20] This suggests that a change in contract structure alters teacher productivity and could in turn affect student performance. My research differs from Jacob's in three key aspects. First, Chicago has a strong teachers' union with collective bargaining power, while North Carolina has no formal teachers' union and collective bargaining is legally prohibited. This difference is reflected in the gap in teacher salaries between Chicago and North Carolina,[21] but not in their teacher tenure policies. Second, my empirical analysis uses student performance as a measure of teacher quality and effectiveness, while Jacob relies on rates of teacher absenteeism as a proxy for teacher productivity. Finally, my research differs from Jacobs' in primary empirical methodology. While he uses a difference-in-difference estimator to examine how tenure status determines teacher behavior, this paper employs a non-parametric specification to discern the relationship between teacher tenure and teacher effectiveness.

Shifting the focus from teacher quality to teacher-student matching, Clotfelter, Ladd, and Vigdor (2006) use North Carolina data to examine how teachers and students are matched, testing the hypothesis that more qualified teachers are more likely to work with students from more affluent backgrounds, with higher levels of parental education, and higher scores on the previous year's state assessment. They find statistically significant evidence of this positive assortative teacher-student matching both between and within schools, with the effect being larger when comparing matching across different schools. This may be attributed to the strong forces of parents and teachers; teachers choose where to work and parents choose where to send their children to school (either directly through a school choice policy, or indirectly by choosing what district to live in). The effect is smaller but still significant within schools, in which teachers and parents have more limited power in influencing teacher-student matching. The authors'

[20] Brian A. Jacob, "The Effect of Employment Protection on Worker Effort: Evidence from Public Schooling," *National Bureau of Economic Research* Working Paper 19655 (2010): 1-40.

[21] National Education Association, *2012-2013 Average Starting Teacher Salaries by State*, compiled from the Collective Bargaining/Member Advocacy Database (December, 2013).

findings also give rise to a debate regarding equity versus efficiency, as they find that the returns from having a higher quality teacher on students' math scores are larger for more affluent, higher performing students. [22] Thus, reassigning higher quality teachers to less affluent, lower performing students could reduce overall average math test scores. This provides a possible explanation for the observed equilibrium of positive assortative teacher-student matching, as well as for the achievement gap between low and high-income students. Though my research is not focused on teacher-student matching, it is still an interesting and relevant topic to consider. A possible extension of my research and a topic for further study could build off of this paper, applying their techniques to an analysis of teacher-student matching for teachers with and without tenure.

IV. Data

All of my empirical research is conducted using data from the NCERDC, a project launched by the Duke University Center for Child and Family Policy and the North Carolina Department of Public Instruction (NCDPI) in 2000.[23] The goal of the project is twofold. First, the NCERDC aims to store and compile existing education data, dating back to the mid-1990s, into a streamlined, standardized format. The project also continues to collect current data on North Carolina's public schools, teachers, and students. In addition to collecting their own data, the NCERDC also has access to data stores from the NCDPI and the National Center for Education Statistics (NCES). Thus, it is a rich source of data for my empirical analysis.

The NCERDC data is broken down into five levels: student, teacher, classroom, school, and district. Each level contains different datasets, which contain various variables that fall under the category of that dataset. Given the focus of my analysis, my research uses data that fall under the student and teacher levels. Though each dataset is compiled over a specific time period, generally from 1996-2011, my research focuses on data collected in 2004, at the conclusion of the 2003-04 school year. This choice of the

---

[22] Charles T. Clotfelter, Helen F. Ladd, and Jacob L. Vigdor, "Teacher-Student Matching and the Assessment of Teacher Effectiveness," *Journal of Human Resources* 41, no. 4 (2006): 778-820.

[23] Duke University Center for Child and Family Policy, North Carolina Education Research Data Center (NCERDC), accessed December 4, 2013, https://childandfamilypolicy.duke.edu/project_detail.php?id=35.

2004 dataset is made to resolve issues of score standardization between years and allow for an adequate range of teacher experience.

The following is a discussion of the primary variables in my analysis, including student test scores, teacher fixed effects, and teacher district experience.

A. End of Grade Tests

Every public school student in North Carolina in grades three through eight is required to take the standardized End of Grade (EOG) exam. Administered at the conclusion of each grade, typically during the last fifteen days of the school year, EOG exams cover English language arts (referred to as 'reading'), mathematics, and science. Each grade is tested in reading and math, with students completing fifth and eighth grade taking an additional science EOG subject test. As the science test is only administered in these two grades, it is omitted from my analysis. A student's raw test score corresponds to one of four levels of mastery in that subject (I-IV). Levels I and II are considered below grade-level, while levels III and IV are considered at or above grade-level. [24] Thus, EOG scores are a natural measure of student performance.

To further introduce the use of EOG test scores, let us look at the test score distributions in 2004 (the relevant year for my analysis). Table 1 is a summary of the EOG test scores for grades four through eight. The table shows the number of observations, minimum and maximum scores, as well as the mean and standard deviation of the score distribution. Since a measure of prior score is necessary for my analysis and EOG tests are first given in grade three, my empirical analysis focuses on grades four through eight. This means that the number of observations in my dataset is reduced by about 16% compared to the original, comprehensive dataset for grades three through eight. To remove outliers and account for students who did not take the exam or those with unreported scores, students with scores of zero in either subject exam are dropped from the dataset. Additionally, as previously explained, all public charter school students are dropped from the dataset as well. These omissions account for only about 5% of the overall dataset, with over 350,000 student observations remaining. Given these trimmings, the findings presented below are for traditional public school students in

[24] NCDPI Division of Accountability Services, *The North Carolina Testing Program: 2012-2013*, July 2012, http://www.ncpublicschools.org/docs/accountability/nctpoverview1213.pdf.

grades four through eight only. These observations are distributed roughly equally between grades, with each of grades four through eight accounting for about 20% of the sample.

Table 1

SUMMARY OF EOG TEST SCORES, GRADES 4-8

READING					
GRADE	TOTAL OBSERVATIONS	MINIMUM SCORE	MAXIMUM SCORE	MEAN	STANDARD DEVIATION
4	68400	224	275	252.8	8.5
5	71471	230	277	257.4	7.8
6	71362	229	283	259.2	8.4
7	71720	228	287	261.7	8.8
8	69819	233	290	264.4	8.7
All (4-8)	352772	224	290	259.1	9.3

MATHEMATICS					
GRADE	TOTAL OBSERVATIONS	MINIMUM SCORE	MAXIMUM SCORE	MEAN	STANDARD DEVIATION
4	68400	232	282	259.6	7.4
5	71471	233	291	263.1	8.7
6	71362	236	293	266.4	9.3
7	71720	231	307	269.1	10.8
8	69819	238	307	272.5	10.6
All (4-8)	352772	231	307	266.2	10.5

SOURCE. – North Carolina Education Research Data Center, EOG Grades 4-8, 2004

Given his or her raw test score, each student is designated a level (I-IV) of mastery in that subject. Table 2 displays the score cutoffs for each of the four levels, by grade, for the 2004 EOG exam. As mentioned above, levels I and II are considered below grade-level, while levels III and IV are considered at or above grade-level. In terms of mastery level, level I is deemed insufficient mastery, level II inconsistent mastery, level III consistent mastery, and level IV superior performance. These score cutoffs represent a developmental scale designed by an in-state psychometrician to reflect learning over time. Since these levels of mastery are determined by the state of North Carolina, they are not meant to serve as a comparison between states. Furthermore, this system is designed so that a student's raw score should increase each year, but the score cutoffs for each grade change so that said student is expected to remain in the same level. For example, a fourth-grade student with a raw score of 244 on the reading exam falls in level III. When that same student is in fifth grade, he or she is expected to score at least 247 to remain in level III. Evidence of this developmental scale is clear in Table 2.

Table 2

EOG SCORE CUTOFFS, GRADES 4-8

READING				
GRADE	LEVEL I	LEVEL II	LEVEL III	LEVEL IV
4	223-235	236-243	244-254	255-275
5	228-238	239-246	247-258	259-277
6	228-241	242-251	252-263	264-283
7	228-242	243-251	252-263	264-287
8	231-243	244-253	254-265	266-290

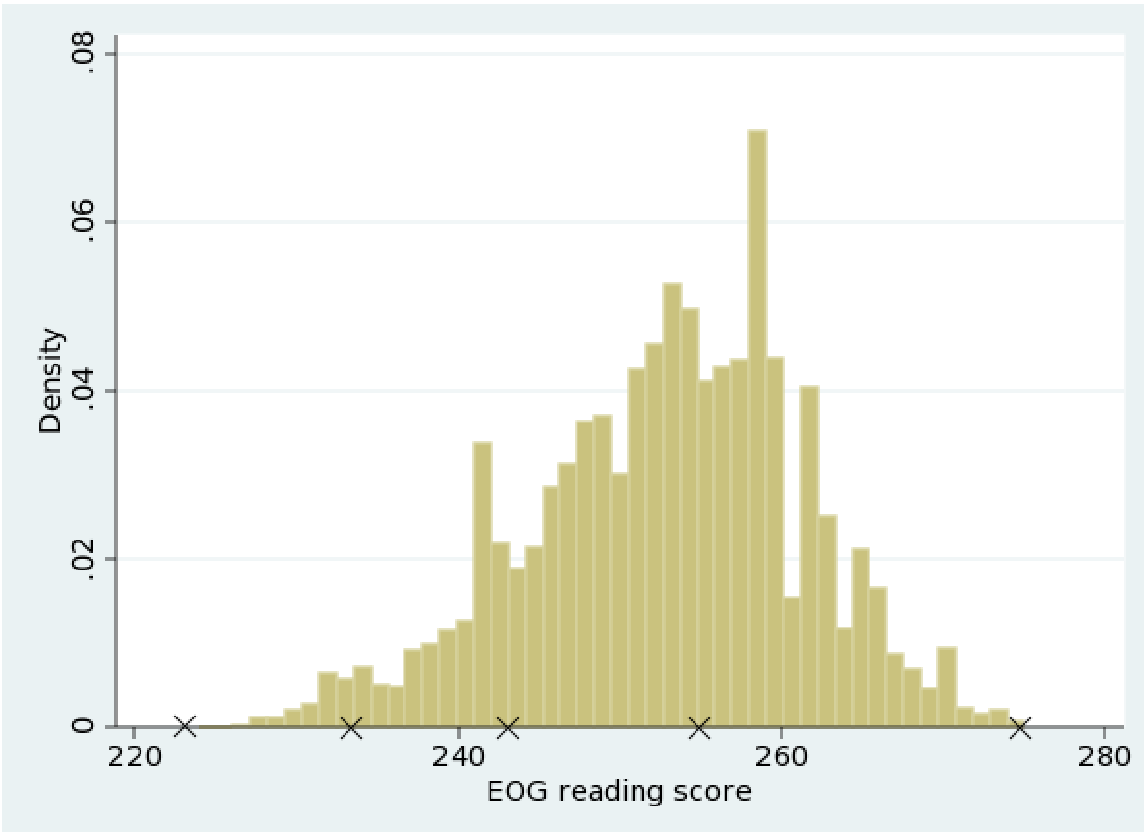
MATHEMATICS				
GRADE	LEVEL I	LEVEL II	LEVEL III	LEVEL IV
4	221-239	240-246	247-257	258-285
5	221-242	243-249	250-259	260-285
6	228-246	247-253	254-264	265-296
7	231-249	250-257	258-266	267-307
8	235-253	254-260	261-271	272-310

SOURCE. – North Carolina Education Research Data Center, Codebook EOG Grades 4-8, 2004

To bring these two tables together in a graphical manner, Figures 1 and 2 display the score distributions for the reading and mathematics exams, respectively, for fourth-grade students in 2004. The x-marks indicate the score cutoffs for each of the four levels, as given in Table 2. The score distributions for grades five through eight look similar and can be found in the Appendix.

Figure 1

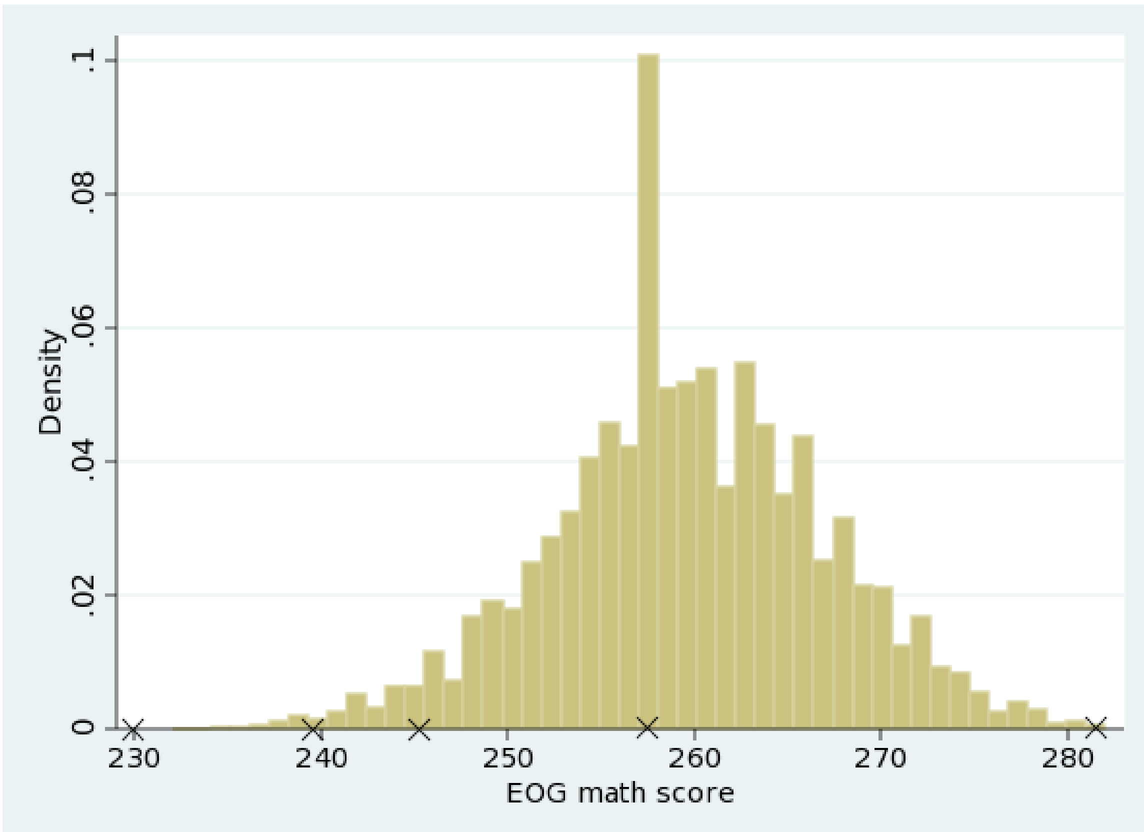
EOG READING SCORE DISTRIBUTION, GRADE 4



SOURCE. – North Carolina Education Research Data Center, EOG Grade 4, 2004

Figure 2

EOG MATHEMATICS SCORE DISTRIBUTION, GRADE 4



SOURCE. – North Carolina Education Research Data Center, EOG Grade 4, 2004

Here, the average reading score (252.8) and the average mathematics score (259.6) fall within the level III and level IV ranges for fourth-grade students, respectively. This indicates that, on average, students performed at or above grade-level and achieved at least a consistent mastery of both subjects. However, the distribution of these scores is highly variant. In both reading and math, a single standard deviation is equal to roughly the range of an entire level. For example, while an average fourth-grade student falls into level III in reading with a score of 252.8, a single standard deviation (8.5) above or below the mean places students in levels II and IV, respectively. This trend is exhibited in both subjects and all grades, suggesting that student test scores are generally a relatively variant measure.

When using student test score performance as a proxy for student achievement, it is also important to note that the mathematics test requires the use of both reading and math skills. Given that the questions are formatted as 'word problems,' a student who has difficulty reading will automatically have difficulty with the mathematics test. Thus, the test scores are not necessarily a pure reflection of solely math knowledge.

Finally, to contextualize the levels of mastery, it is important to keep in mind the policy implications. From a program and policy perspective, most teacher efforts and interventions at any level (classroom, school, district, state, national) are aimed at level II and III students. This is unsurprising given the monetary incentives attached to school accountability. Neal and Schanzenbach (2010) find concrete evidence of this phenomenon after the introduction of No Child Left Behind in 2002.[25] For example, if a school needs a certain percentage of students to perform at grade-level in the upcoming year in order to receive increased funding, teachers and schools tend to focus their attention on getting students who are near the threshold up to grade-level (level II students) and making sure students right at the threshold do not fall below it (level III students). This demonstrates the importance, particularly from a policy standpoint, of small, incremental changes in student standardized test score performance.

B. Teacher Fixed Effects

In order to link student EOG test score performance to teacher quality, I construct a measure called teacher fixed effects. This variable estimates a teacher's average effect on the test scores of his or her students in a given year. To construct this variable, I first sort the EOG score datasets by student. Since this dataset is at the student level, each set of test scores in a given year is linked to a specific student by an identification code (*mastid*). I also introduce a combined score measure (*comb = read + math*), which is created simply by summing a student's scores on the reading and mathematics EOG exams. By merging the datasets from the previous year, I am able to create prior score variables for each test ('*readprior*' and '*mathprior*', as well as '*combprior*'). For

---

[25] Derek Neal and Diane W. Schanzenbach, "Left Behind By Design: Proficiency Counts and Test-Based Accountability," *Review of Economics and Statistics* 92 (2010): 263-283.

example, a 2004 fourth-grade student's prior EOG scores are given in the 2003 EOG dataset for grade three. By merging these datasets together and sorting by student identification number, I am able to construct the prior score variables. These same steps are taken to construct prior score measures for students in grades four through eight (grade three is omitted as there is no EOG exam for grade two, and thus no measure of prior score) in 2004. Using this data, I run a regression comparing student EOG test scores in 2004 to their scores in 2003. In order to allow for grade-specific intercepts, I include an indicator variable for grades four through seven (the indicator for grade eight is omitted to prevent collinearity). To account for varying slopes across grades, I include interaction terms between the grade-specific indicators and the prior score variable.

The equations below were used to specify the regressions:

$$(1) \quad read_i = a_0 + a_{1g} G_{ig} + a_{2g}(G_{ig} \times readprior_i) + u_i,$$

$$(2) \quad math_i = a_0 + a_{1g} G_{ig} + a_{2g}(G_{ig} \times mathprior_i) + u_i,$$

$$(3) \quad comb_i = a_0 + a_{1g} G_{ig} + a_{2g}(G_{ig} \times combprior_i) + u_i,$$

where $read_i$ is student i's EOG reading score in the current year (2004), G_{ig} is an indicator variable for student i's grade (g) in the current year, and G_{ig} x $readprior_i$ is the interaction term between student i's grade in the current year and his or her EOG reading score in the previous year (2003). The math and combined score variables are specified in an identical fashion using the EOG math and combined test scores. Table 3 displays the results of these regressions.

Table 3

REGRESSION OF CURRENT EOG SCORES ON PRIOR EOG SCORES, GRADES 4-8

READING			
INDEPENDENT VARIABLE	COEFFICIENT	STANDARD ERROR	P-VALUE
Grade Indicators (G_{ig})			
Grade 4 (G_4)	7.0	.79	0.00
Grade 5 (G_5)	19.5	.80	0.00
Grade 6 (G_6)	-15.6	.84	0.00
Grade 7 (G_7)	-17.1	.82	0.00

Constant (a_0)	54.3	.57	0.00
Interaction Terms (*G_{ig} x readprior$_i$*)			
G_4 x readprior	.77	.002	0.00
G_5 x readprior	.73	.002	0.00
G_6 x readprior	.86	.002	0.00
G_7 x readprior	.87	.002	0.00
G_8 x readprior	.80	.002	0.00
			Adjusted R^2 = .7211

MATHEMATICS

INDEPENDENT VARIABLE	COEFFICIENT	STANDARD ERROR	P-VALUE
Grade Indicators (*G_{ig}*)			
Grade 4 (G_4)	-39.1	.92	0.00
Grade 5 (G_5)	-41.2	.80	0.00
Grade 6 (G_6)	-13.7	.73	0.00
Grade 7 (G_7)	-48.2	.72	0.00
Constant (a_0)	42.2	.47	0.00
Interaction Terms (*G_{ig} x mathprior$_i$*)			
G_4 x mathprior	1.01	.003	0.00
G_5 x mathprior	1.01	.003	0.00
G_6 x mathprior	.91	.002	0.00
G_7 x mathprior	1.03	.002	0.00
G_8 x mathprior	.86	.002	0.00
			Adjusted R^2 = .7910

COMBINED

INDEPENDENT	COEFFICIENT	STANDARD ERROR	P-VALUE

VARIABLE			
Grade Indicators (G_{ig})			
Grade 4 (G_4)	-21.3	1.40	0.00
Grade 5 (G_5)	-16.8	1.33	0.00
Grade 6 (G_6)	-31.2	1.31	0.00
Grade 7 (G_7)	-68.6	1.28	0.00
Constant (a_0)	69.5	.86	0.00
Interaction Terms (G_{ig} x *combprior*$_i$)			
G_4 x combprior	.92	.002	0.00
G_5 x combprior	.91	.002	0.00
G_6 x combprior	.94	.002	0.00
G_7 x combprior	1.00	.002	0.00
G_8 x combprior	.88	.002	0.00
			Adjusted R^2 = .8332

SOURCE. – North Carolina Education Research Data Center, EOG Grades 4-8, 2003-2004

From these regressions, it is clear that prior EOG scores are a strong predictor of current student test performance. For each grade, the coefficients on both the reading and math prior score interaction terms are highly statistically significant (p-value= 0.00) and positive, indicating, as expected, that high test scores in the previous year are correlated with high scores in the current year. For example, this regression predicts that on the fourth grade EOG reading exam in 2004, a given student will score 77% of his or her reported score on the third grade EOG reading exam in 2003, plus a grade-specific constant of 61.3 points ($G_4 + a_0$). Given the shift in mastery level score cutoffs between grades, this translates to roughly the same level of mastery between years. Thus, a student who performed at grade-level (level III) in grade three is expected to again score at grade-level for grade four in the following year. This is consistent with the previously discussed developmental scale structure of the EOG exams.

In interpreting these results, it is essential to note that the relationship between current scores and prior scores represents correlation, not causation. The prior score variable absorbs the compounding effects of individual student characteristics (sex, ethnicity, parental education level, ability, etc.) and teacher differences. This explains why, in all of the above regressions, the R^2 values are relatively high. When all grades are aggregated together, prior test scores explain roughly 72% of the variation in current EOG reading scores, 79% of the variation in current EOG math scores, and 83% of the variation in current EOG combined scores. Thus, a natural next step is to separate out the effects of student characteristics and teacher differences.

Teachers and students are linked in the data through an identification coding system. In the EOG dataset, each student is linked to the teacher (through the *teachid* variable) that proctored his or her exam. For the elementary grades (grades three through five), the teacher who proctors a student's exam is almost always the student's classroom teacher. Though this match rate is lower for middle school students (grades six through eight), for the sake of my analysis (which focuses on grades four through eight) I assume that the specified teacher is the student's actual classroom teacher.

Ideally, to isolate teacher effects, I would like to add an indicator variable for each individual teacher into regressions (1), (2), and (3). As this would need to include a variable for each individual teacher, it is obviously impossible given the constraints of the data analysis software. Instead, using the same dataset, I use an absorbing indicators regression to estimate teacher fixed effects for each of the two EOG subject scores. This absorbing indicators regression is exactly equivalent to a standard regression with an indicator variable for each teacher.

Table 4 displays the results of these three regressions, specified below.

(4) $read_i = a_0 + a_{1g}\, G_{ig} + a_{2g}(G_{ig}\ x\ readprior_i) + TeacherFE_j + u_i,$

(5) $math_i = a_0 + a_{1g}\, G_{ig} + a_{2g}(G_{ig}\ x\ mathprior_i) + TeacherFE_j + u_i,$

(6) $comb_i = a_0 + a_{1g}\, G_{ig} + a_{2g}(G_{ig}\ x\ combprior_i) + TeacherFE_j + u_i,$

where $TeacherFE_j$ is teacher j's estimated fixed effect on student i's current (2004) EOG score.

Table 4

ABSORBED LINEAR REGRESSION OF CURRENT EOG SCORES ON PRIOR EOG SCORES, GRADES 4-8

READING			
INDEPENDENT VARIABLE	COEFFICIENT	STANDARD ERROR	P-VALUE
Grade Indicators (G_{ig})			
Grade 4 (G_4)	-5.4	1.17	0.00
Grade 5 (G_5)	7.8	1.16	0.00
Grade 6 (G_6)	-16.3	1.01	0.00
Grade 7 (G_7)	-17.0	.98	0.00
Constant (a_0)	74.1	.74	0.00
Interaction Terms (G_{ig} x readprior$_i$)			
G_4 x readprior	.74	.002	0.00
G_5 x readprior	.69	.002	0.00
G_6 x readprior	.79	.003	0.00
G_7 x readprior	.80	.003	0.00
G_8 x readprior	.73	.003	0.00
			Adjusted R^2 = .7400

MATHEMATICS			
INDEPENDENT VARIABLE	COEFFICIENT	STANDARD ERROR	P-VALUE
Grade Indicators (G_{ig})			
Grade 4 (G_4)	-46.8	1.22	0.00
Grade 5 (G_5)	-47.5	1.12	0.00
Grade 6 (G_6)	-12.6	.86	0.00

Grade 7 (G_7)	-42.4	.85	0.00
Constant (a_0)	57.7	.63	0.00
Interaction Terms (G_{ig} x *mathprior*$_i$)			
G_4 x mathprior	.97	.003	0.00
G_5 x mathprior	.97	.003	0.00
G_6 x mathprior	.85	.002	0.00
G_7 x mathprior	.96	.002	0.00
G_8 x mathprior	.80	.002	0.00
			Adjusted R^2 = .8237

COMBINED

INDEPENDENT VARIABLE	COEFFICIENT	STANDARD ERROR	P-VALUE
Grade Indicators (G_{ig})			
Grade 4 (G_4)	-35.5	1.95	0.00
Grade 5 (G_5)	-28.9	1.88	0.00
Grade 6 (G_6)	-30.3	1.60	0.00
Grade 7 (G_7)	-62.6	1.56	0.00
Constant (a_0)	96.2	1.16	0.00
Interaction Terms (G_{ig} x *combprior*$_i$)			
G_4 x combprior	.89	.002	0.00
G_5 x combprior	.88	.002	0.00
G_6 x combprior	.89	.002	0.00
G_7 x combprior	.95	.002	0.00
G_8 x combprior	.83	.002	0.00
			Adjusted R^2 = .8529

SOURCE. – North Carolina Education Research Data Center, EOG Grades 4-8, 2003-2004

As expected, the coefficients on the prior reading, math, and combined scores in Table 4 fell (as compared to regressions (1), (2), and (3), respectively). Though there is still a significant positive correlation between current score and prior score, the magnitude of this relationship is smaller when teacher effects are separated from prior scores. Let us return to our previous example. As previously noted, regression (1) in Table 3 predicts that for a fourth-grade student in 2004, his or her EOG reading score will be 77% of his or her score on the EOG reading exam for grade three in 2003 (plus a constant of 61.3 points). With the addition of teacher fixed effects, regression (3) in Table 4 now predicts that the student's score on the EOG reading exam in 2004 will be 74% of his or her prior reading score, *plus* the fixed effect of his or her teacher in 2004 (plus a constant of 68.7 points). This demonstrates that teachers do have an impact on student test scores.

The adjusted R^2 values in Table 4 indicate that prior scores and teacher fixed effects together account for around 74% of the variation in current EOG reading test scores, 82% of the variation in current EOG math test scores, and 85% of the variation in current EOG combined test scores. This increase in adjusted R^2 (from 72%, 79%, and 83%, respectively) is expected, given that regressions (4), (5), and (6) introduce a separate variable for teacher differences into the equation.

Let us now turn to a discussion of teacher fixed effects, a proxy for teacher quality. Teacher fixed effects are defined as a teacher's average effect on the test scores of his or her students in a given year. For example, let's say that a teacher's fixed effect on EOG reading scores is measured to be 1.5. This means that averaging across all students in his or her classroom, a 1.5 point increase in reading test scores is attributable to the teacher in that given year. This is meant to capture the aspects of teaching that are difficult to quantify and observe, but are constant within a classroom.

Using the fixed effects estimation given by the absorbed linear regressions above, Table 5 gives a statistical summary of teacher fixed effects on EOG scores. Note that these results are consistent with prior literature, which suggests that teachers are responsible for roughly 8-15% of a standard deviation in student test scores in a given year.[26] Figures 3, 4, and 5 provide a graphical distribution of the teacher fixed effects on each subject test score.

Table 5

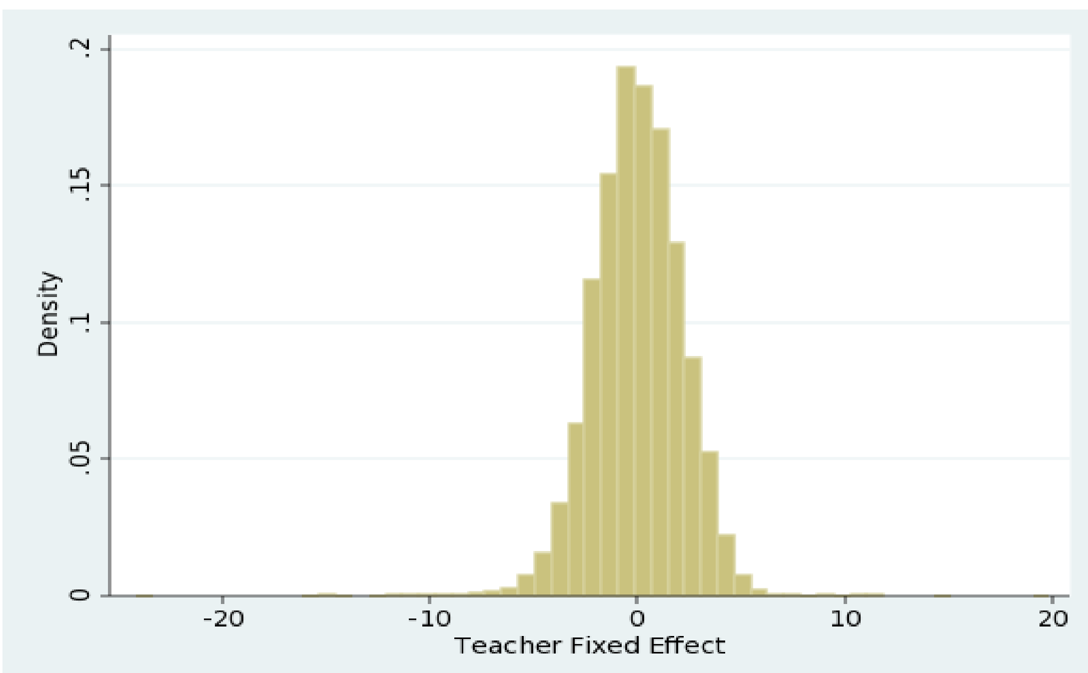
SUMMARY OF TEACHER FIXED EFFECTS ON EOG SCORES, GRADES 4-8

SUBJECT	TOTAL OBSERVATIONS	MINIMUM	MAXIMUM	MEAN	STANDARD DEVIATION
Reading	328987	-24.2	19.9	$8.43e^{-10}$	2.1
Mathematics	328987	-19.5	15.9	$2.99e^{-10}$	2.4
Combined	328987	-26.6	33.6	$3.39e^{-10}$	3.9

SOURCE. – North Carolina Education Research Data Center, EOG Grades 4-8, 2003-2004

Figure 3

DISTRIBUTION OF TEACHER FIXED EFFECTS ON EOG READING SCORES, GRADES 4-8

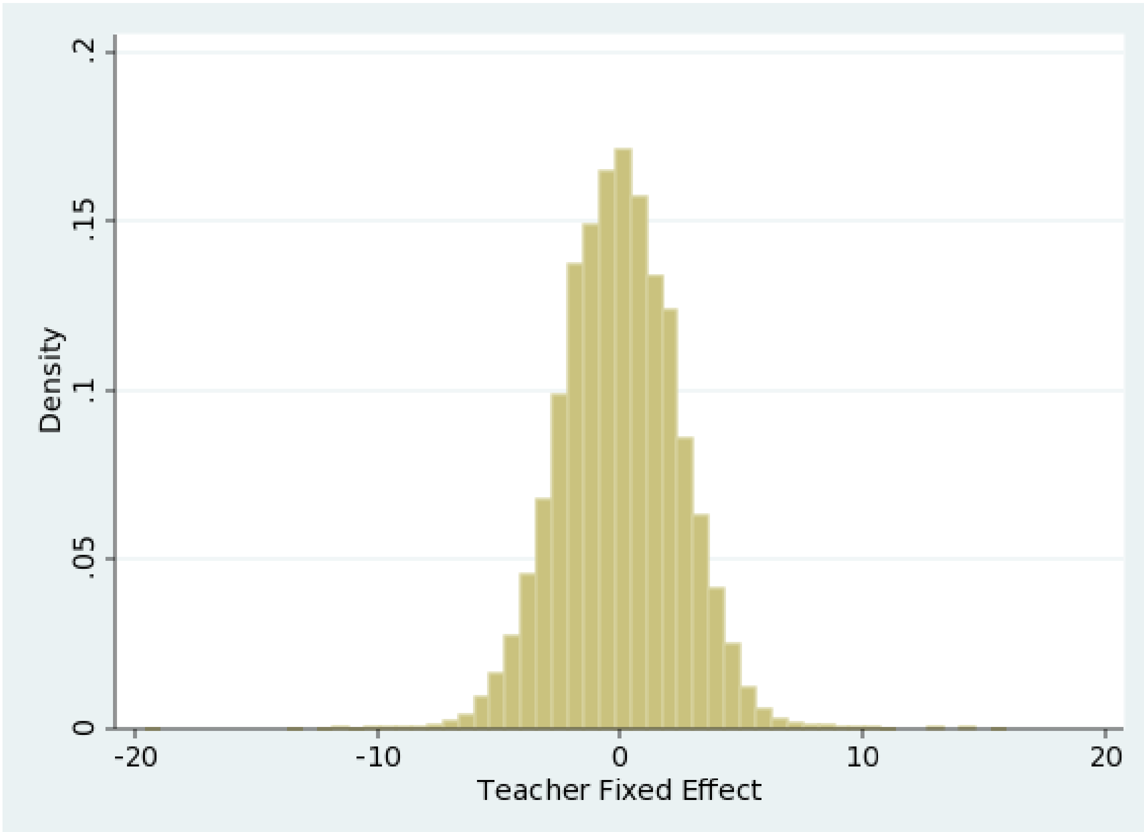


SOURCE. – North Carolina Education Research Data Center, EOG Grades 4-8, 2003-2004

[26] Hugh Macartney, "The Dynamic Effects of Educational Accountability," *National Bureau of Economic Research* Working Paper 19915 (2014): 4.

Figure 4

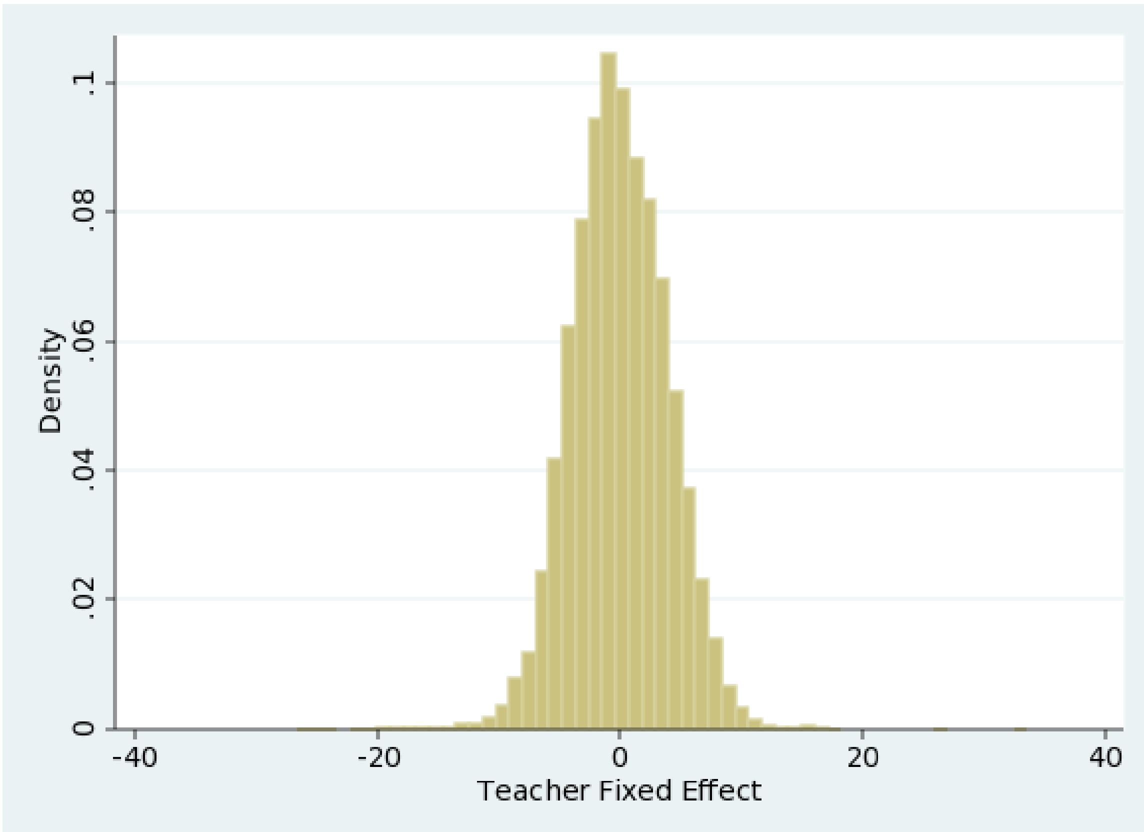
DISTRIBUTION OF TEACHER FIXED EFFECTS ON EOG MATHEMATICS SCORES, GRADES 4-8



SOURCE. – North Carolina Education Research Data Center, EOG Grades 4-8, 2003-2004

Figure 5

DISTRIBUTION OF TEACHER FIXED EFFECTS ON EOG COMBINED SCORES, GRADES 4-8



SOURCE. – North Carolina Education Research Data Center, EOG Grades 4-8, 2003-2004

For reading, mathematics, and combined scores, teacher fixed effects are normally distributed around a mean of roughly zero. This is the case by design, as teacher fixed effects are constructed by taking the residuals from a standard regression without teacher fixed effects (which are normally distributed around a mean of zero) and taking means by teacher. Since there are a sufficient number of teachers, the distribution remains normal. However, when disaggregating teacher fixed effects by years of teacher district experience, a different pattern emerges. This is addressed in Section V.

C. Teacher District Experience

The final explanatory variable central to this analysis is teacher district experience, defined as consecutive years of teaching experience in the same district. As tenure is awarded to teachers by district (see section II for more detail), it is necessary to separate years teaching within the same district from total years of teaching experience. To construct this district experience variable, I first sort the teacher pay dataset (which contains teacher experience) by teacher identification code (*teachid*). Given the range of the previously specified tenure bubble (years one through eight), I begin with the 2004 dataset to allow for up to eight years of experience within a district (collection of this dataset began in 1996). Using a district code (*leacode)*, the data indicates which school district employs each teacher in that given year. By merging the data for each year and looking at the matching between the district code in a given year with that in the previous year, I am able to build a district experience variable for each teacher, ranging from one to eight years of teaching experience within a district. To clarify, a teacher with one year of district experience has just completed his or her first year of teaching within a district, while a teacher with eight years of district experience has taught in the same district for eight or more consecutive years. This district experience variable spans the range of the previously specified tenure bubble. It should be noted that this variable measures district-specific experience, not necessarily experience teaching in the same school or at the same grade. Additionally, as teachers can move between districts, it does not necessarily capture total years of teaching experience either. However, since tenure is awarded at the district level, district experience is sufficient for this analysis.

As previously explained in the background section, teachers in North Carolina are eligible for tenure after four years of teaching in the same district. Thus, it is reasonable to assume that a teacher with more than four years of experience in a district has tenure, while those with four or fewer years of district experience are probationary (non-tenured) teachers. Using this fact along with the district experience variable, I construct an indicator variable for whether or not a teacher has tenure (tenure=1 if teacher is tenured, tenure=0 otherwise). Table 6 provides a summary of the district experience and tenure variables.

Table 6

YEARS OF TEACHER EXPERIENCE WITHIN A SCHOOL DISTRICT

VARIABLE	TOTAL OBSERVATIONS	MEAN	STANDARD DEVIATION	MEDIAN

DISTRICT EXPERIENCE	352772	5.9	2.3	7.0
TENURE INDICATOR	352772	.67	.47	1.0

SOURCE. – North Carolina Education Research Data Center, Teacher Pay, 1996-2004

The average teacher in this sample has roughly 5.9 years of experience in a specific district, indicating that the average teacher is tenured. This sample exhibits a large standard deviation relative to its mean, thus the median (seven years of district experience) is presented as an alternative measure. These results are interpreted keeping in mind that a teacher with a district experience value of eight years actually has eight *or more* years of experience within a district, thus actual average district experience is potentially higher. However, given the focus of my analysis, this does not impact my empirical specification or results.

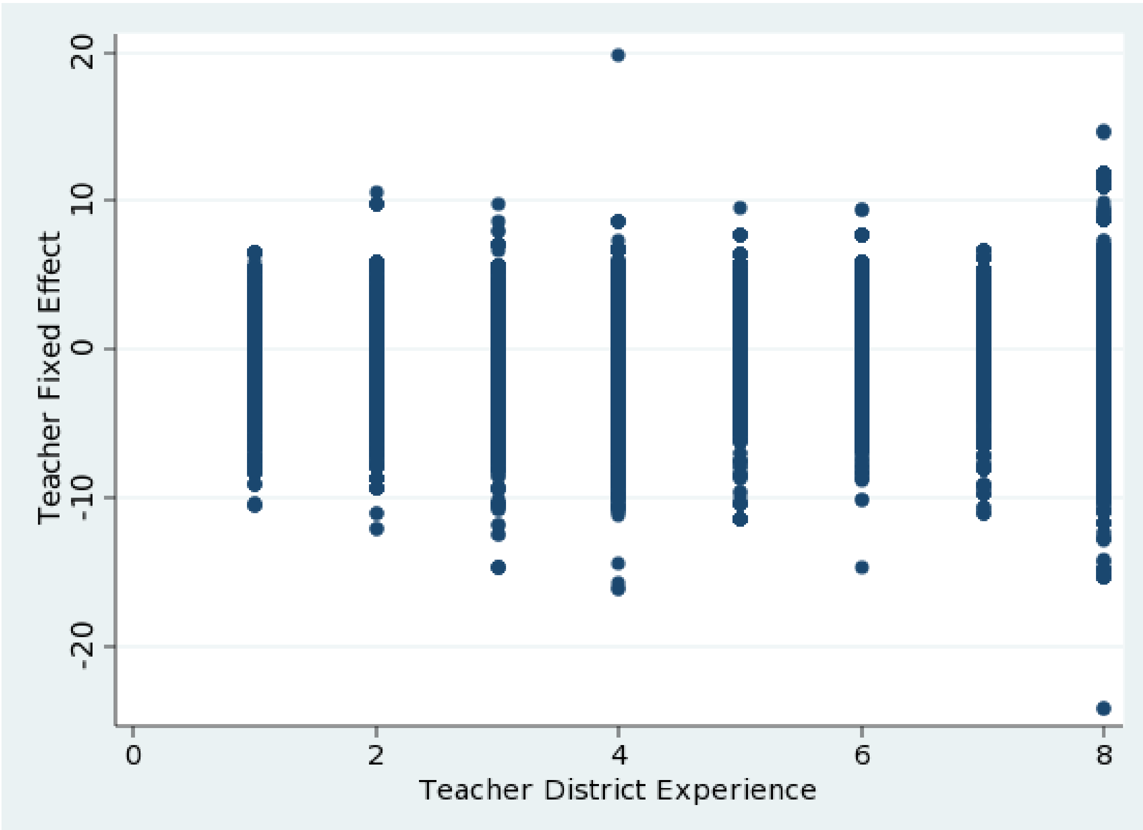
From the tenure indicator variable, we can conclude that 67% of teachers in this sample have tenure. Though this is interesting in and of itself, the relationship of these two variables (district experience and tenure status) to teacher quality is essential to my main empirical analysis.

V. Empirical Specification and Results

Given the discussed measures of teacher tenure (given by district experience) and teacher quality (given by teacher fixed effects), the discussion now moves to an examination of the relationship between the two. Since this relationship is unknown, a non-parametric specification is used to compare average teacher quality across the tenure bubble (district experience years one through eight). The first step in this analysis is a plot comparing district experience (x-axis) and teacher fixed effects (y-axis). Each point represents a single teacher's fixed effect measure relative to his or her years of district experience. Figures 6, 7, and 8 display this plot for teacher fixed effects on EOG reading, mathematics, and combined scores, respectively.

Figure 6

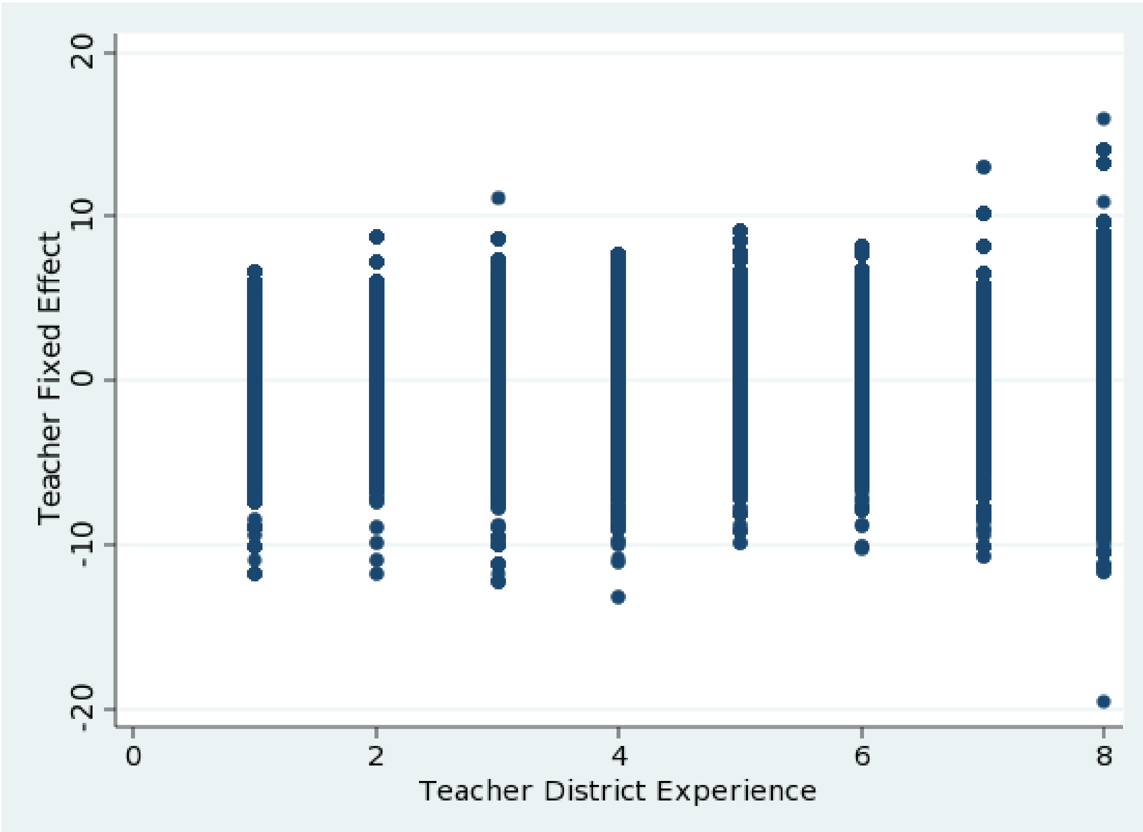
TEACHER FIXED EFFECTS ON EOG READING SCORES BY TEACHER DISTRICT EXPERIENCE



SOURCE. – North Carolina Education Research Data Center, Teacher Pay, 1996-2004

Figure 7

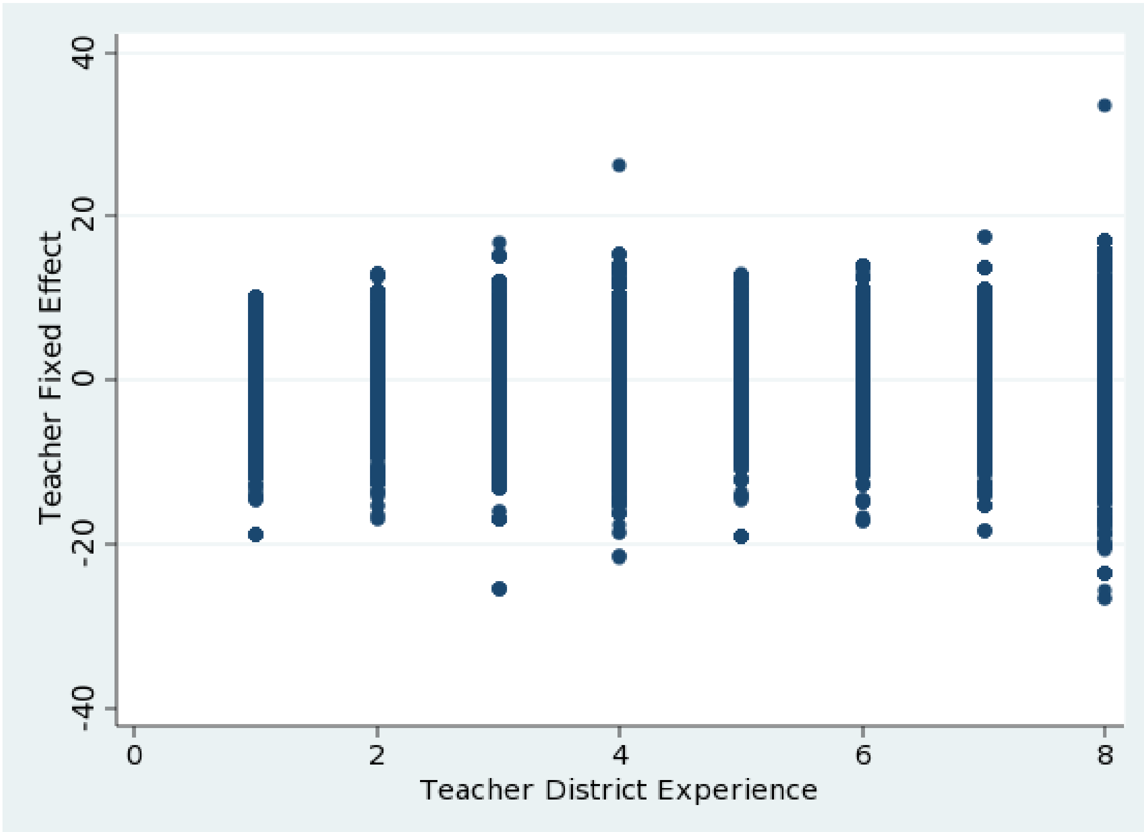
TEACHER FIXED EFFECTS ON EOG MATHEMATICS SCORES BY TEACHER DISTRICT EXPERIENCE



SOURCE. – North Carolina Education Research Data Center, Teacher Pay, 1996-2004

Figure 8

TEACHER FIXED EFFECTS ON EOG COMBINED SCORES BY TEACHER DISTRICT EXPERIENCE



SOURCE. – North Carolina Education Research Data Center, Teacher Pay, 1996-2004

Each point on Figures 6, 7, and 8 represents a single teacher's unique *(years of district experience, fixed effect)* pair. As depicted in each of the figures, these points come together to create a vertical line for each year of district experience. The length of this line represents the range of teacher fixed effects for teachers within that year of district experience (the line for the eighth year represents teachers with eight or more years of district experience). In other words, the length of each line is the range of teacher quality within that year of district experience. One can also imagine that these three figures have a third dimension, which is the frequency of observations along any given vertical line. The third dimension of each vertical line gives the distribution of teacher fixed effects within each year of district experience.

Since this portion of the analysis is interested in teachers within a given level of district experience in the aggregate, the fixed effects are then averaged. Taking an average of the teacher fixed effects within each year gives a single point for each district experience year. This gives eight pairs of (*year of district experience*, *average teacher fixed effect*) data points for each subject. These data points, as well as the confidence intervals around the mean, are given in Table 7, for fixed effects on reading, mathematics, and combined scores. Figures 9, 10, and 11 display this information graphically. The dots surrounding the mean on the figures represent the upper and lower bounds of the 95% confidence intervals given in Table 7.

Table 7

AVERAGE TEACHER FIXED EFFECTS BY YEARS OF DISTRICT EXPERIENCE

READING				
YEARS OF DISTRICT EXPERIENCE	TOTAL OBSERVATIONS	MEAN	STANDARD DEVIATION	CONFIDENCE INTERVAL (95%)
1	13772	-.36	2.08	[-.40, -.33]
2	15374	-.21	2.12	[-.24, -.18]
3	36035	-.30	2.08	[-.32, -.28]
4	41652	-.30	2.11	[-.32, -.28]
5	25718	-.09	2.07	[-.11, -.06]
6	22223	.04	2.04	[.01, .07]
7	18185	.11	2.03	[.08, .13]
8	156028	.20	2.06	[.19, .21]

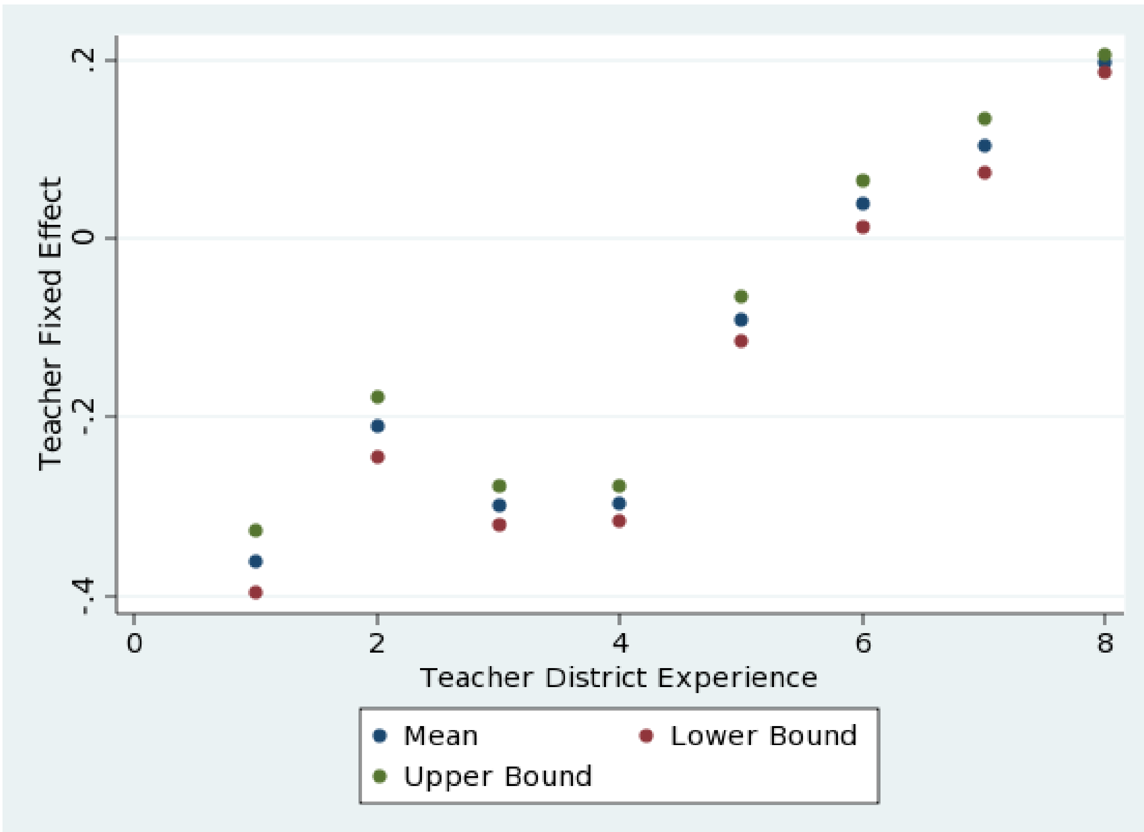
Mathematics				
Years of District Experience	Total Observations	Mean	Standard Deviation	Confidence Interval (95%)
1	13772	-.32	2.42	[-.36, -.28]
2	15374	-.25	2.38	[-.29, -.22]
3	36035	-.42	2.36	[-.44, -.39]
4	41652	-.31	2.25	[-.33, -.29]
5	25718	-.07	2.37	[-.10, -.04]
6	22223	.04	2.31	[.01, .07]
7	18185	.06	2.29	[.03, .10]
8	156028	.23	2.36	[.22, .24]

Combined				
Years of District Experience	Total Observations	Mean	Standard Deviation	Confidence Interval (95%)
1	13772	-.57	3.88	[-.63, -.50]
2	15374	-.37	3.87	[-.43, -.30]
3	36035	-.64	3.83	[-.68, -.60]
4	41652	-.53	3.81	[-.56, -.49]
5	25718	-.13	3.86	[-.18, -.09]
6	22223	.06	3.81	[.01, .11]
7	18185	.12	3.79	[.06, .17]
8	156028	.38	3.85	[.36, .39]

Source. – North Carolina Education Research Data Center, Teacher Pay, 1996-2004

Figure 9

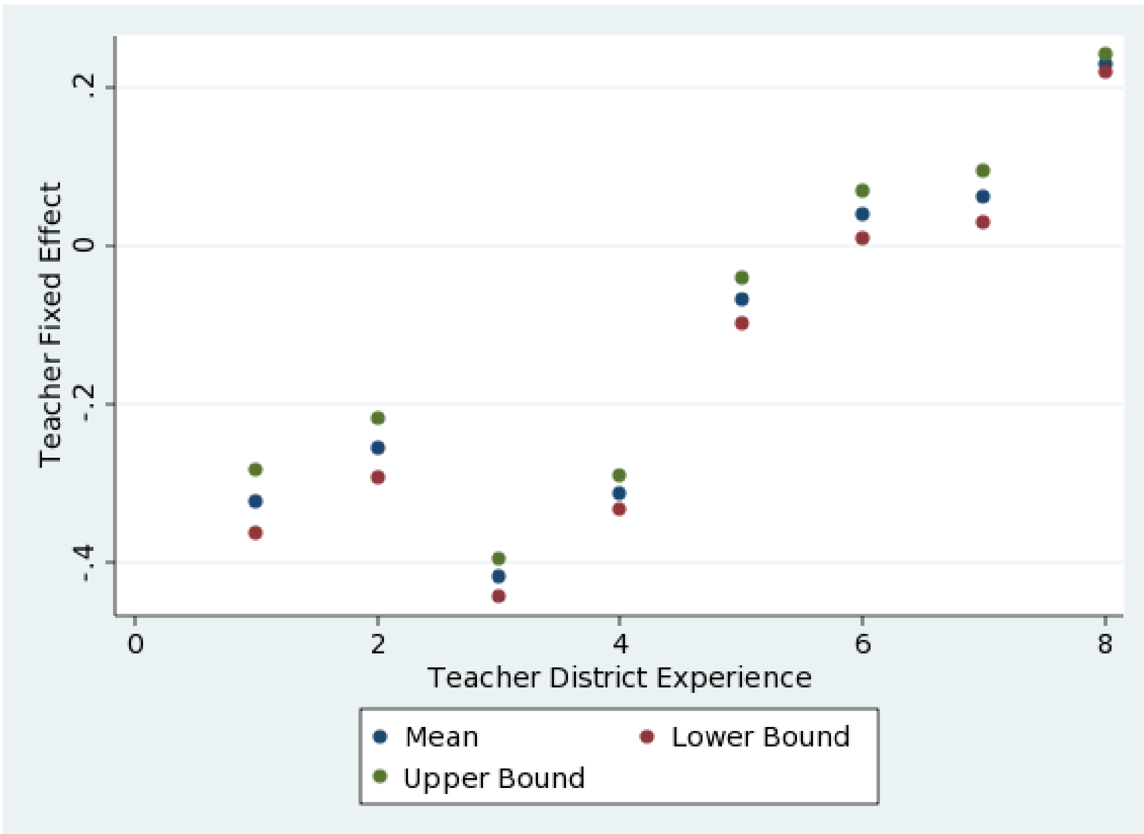
AVERAGE TEACHER FIXED EFFECTS ON EOG READING SCORES BY YEARS OF DISTRICT EXPERIENCE



SOURCE. – North Carolina Education Research Data Center, Teacher Pay, 1996-2004

Figure 10

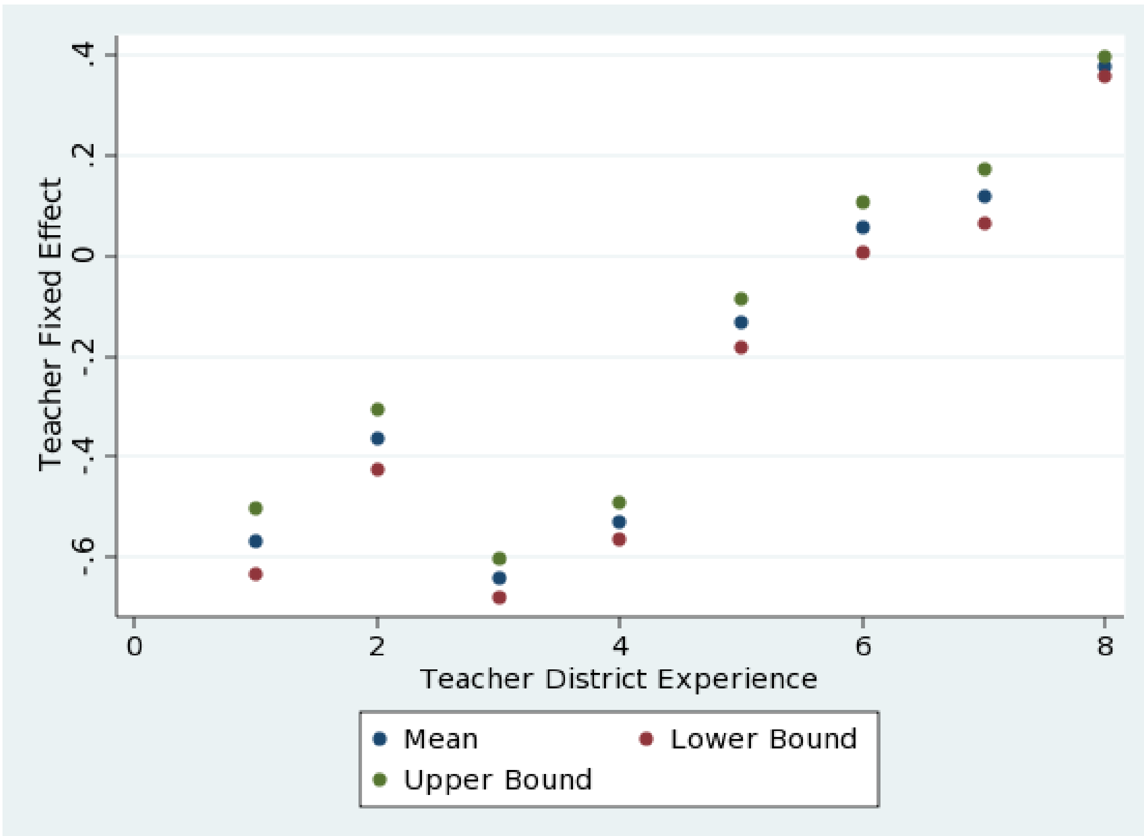
AVERAGE TEACHER FIXED EFFECTS ON EOG MATHEMATICS SCORES BY YEARS OF DISTRICT EXPERIENCE



SOURCE. – North Carolina Education Research Data Center, Teacher Pay, 1996-2004

Figure 11

AVERAGE TEACHER FIXED EFFECTS ON EOG COMBINED SCORES BY YEARS OF DISTRICT EXPERIENCE



SOURCE. – North Carolina Education Research Data Center, Teacher Pay, 1996-2004

As expected, these plots show there is an observed positive correlation between average teacher fixed effects and years of district experience. This means that, in general, average teacher quality increases as years of district experience increase. For example, on average, teachers with four years of district teaching experience are more effective than those with three years of district teaching experience. This comparison is not necessarily between individual teachers in the same district, but rather in the aggregate of all teachers with a given level of district experience, across districts.

Furthermore, it is clear that there is an empirically observable relationship between years of district experience and teacher fixed effects across the tenure bubble.

Given that teachers in North Carolina are eligible for tenure after four years of teaching in the same district, an analysis of tenure policy draws our attention to the change in average teacher fixed effects between years four and five. For fixed effects on EOG reading scores (Figure 9), there is a large jump in the graph in this interval. Consistent with the previous discussion, teachers with five years of district experience are on average of higher quality than those with four years of district experience. However, this jump is a clear break in the trend of the plot. This picture suggests that there is a change in teacher quality between four and five years of district experience beyond just the expected increase that comes with an additional year of teaching experience. This means that low quality teachers do not persist in the same district after year four, causing the average fixed effect for teachers with five years of district experience to be substantially higher. The teacher fixed effects on EOG math scores (Figure 10) and combined scores (Figure 11) exhibit a similar trend.

Looking at the confidence intervals surrounding in the means on each figure, it is clear that this effect is statistically significant at the 5% level. This can be seen by the fact that the mean within year four is not included in the 95% confidence interval for year five. It should be noted that, in each of the three plots, the confidence interval for year eight is very narrow. Given that year eight includes all teachers with eight or more years of experience in a district, it is unsurprising that this category includes more observations, and thus exhibits a very narrow confidence interval around the mean.

In the context of tenure, these results suggest that the tenure process exhibits strong positive selection, meaning that low quality teachers are not awarded tenure. This is consistent with the goals of teacher hiring policy- to retain high quality teachers while removing those that are less effective. Thus, in this sense, North Carolina's current teacher tenure policy proves to be effective.

This analysis demonstrates that not only is there a sharp break in the figures at the tenure bubble, but also suggests that returns to district level experience on teaching effectiveness increase after the tenure bubble. Though less pronounced, this trend is visible in each of the three figures and is statistically significant at the 5% level. This result suggests that the tenure process not only selects the more effective teachers at the time, but also selects teachers with a higher propensity for growth. An alternative argument, which is used to justify tenure at the research and university level, is that the job security provided by tenure encourages teacher innovation and improvement. While further analysis is needed to come to a decisive conclusion, it is likely that this result is the product of a combination of these two explanations.

VI. Conclusion

As there are other possible explanations for the observed relationship between teacher tenure and teacher quality, this analysis does not provide a definitive answer on its causation. However, it is clear that on average, tenured teachers are of higher demonstrated quality than probationary teachers, beyond the expected increase in quality attributable to more experienced teachers**.** Thus, it is clear that teacher tenure policy does play an important role in raising overall teacher quality and effectiveness.

This study of tenure is particularly timely, as the North Carolina legislature recently passed a sweeping tenure reform, which is slated to gradually go into effect over the next five years.[27] The new tenure reform eliminates the existing four-year teacher tenure eligibility rule, limiting tenure only to the 25% percent of public school teachers in the state who are deemed most effective (effectiveness criteria has not yet been specified). Other teachers will be eligible for four-, two-, and one-year contracts, depending on their experience and demonstrated effectiveness. North Carolina is not the first state to radically alter its tenure policies, as Louisiana, South Dakota, Idaho, Colorado, and Florida, among others, have recently made changes as well.[28]

Though the North Carolina legislation has already been passed, it has not yet been implemented. Thus, it is important to examine the current tenure policy and its relationship to teacher quality before any drastic changes are made. My research serves two unique purposes in this respect. Qualitatively, it provides a summary of the existing policy and its implementation. Empirically, the introduction of the tenure indicator variable and tenure bubble is unique to my analysis and adds another component to the already rich literature on teacher quality. Furthermore, by focusing on the fact that teacher tenure is awarded at the district level, rather than on total years of teaching

---

[27] Bonner, "Proposed NC Budget Would End Teacher Tenure, Pay Tuition Vouchers," July 21, 2013,

[28] Adrienne Lu, "North Carolina Ends Teacher Tenure," *Stateline*, July 29, 2013, http://www.pewstates.org/projects/stateline/headlines/north-carolina-ends-teacher-tenure-85899493655.

experience, my research deviates from the existing literature examining the relationship between teacher experience and teacher effectiveness.

This empirical analysis could inform policy research as well, potentially revealing any unanticipated effects a change in tenure policy may have on teacher effectiveness. Once changes are made to North Carolina's teacher tenure policy, it would be interesting to do follow up research in the future examining how teacher effectiveness changes after the implementation of the new policy.

It should also be noted that the analysis and results presented here apply to a non-union state. As an area for further research, it would be interesting to see if the same findings are present in states where there is a strong teachers' union. One could also imagine carrying out this analysis in a state or district that explicitly uses student test score performance to dictate teacher hiring and retention practices.

Bibliography

Bonner, Lynn. "Proposed NC Budget Would End Teacher Tenure, Pay Tuition Vouchers." *News Observer,* July 21, 2013. http://www.newsobserver.com/2013/07/21/3046589/state-budget-would-end-teacher.html.

Bonner, Lynn and Jane Stancil. "McCrory Seeks to Raise Teachers Starting Pay." *News Observer*, February 10, 2014. http://www.newsobserver.com/2014/02/10/3608266/mccrory-increase-teachers-base.html.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. "The Long-Term Impacts of Teachers: Teacher Value-Added and Student Outcomes in Adulthood." *National Bureau of Economic Research* Working Paper 19423 (2013): 1-42.

Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. "Teacher-Student Matching and the Assessment of Teacher Effective." *Journal of Human Resources* 41, no. 4 (2006): 778-820.

Duke University Center for Child and Family Policy. North Carolina Education Research Data Center (NCERDC). Accessed December 4, 2013. https://childandfamilypolicy.duke.edu/project_detail.php?id=35.

Harris, Douglas N. and Tim R. Sass. "Teacher Training, Teacher Quality, and Student Achievement." *Journal of Public Economics* 95 (2011): 798-812.

Haskins, Ron and Susanna Loeb. "A Plan to Improve the Quality of Teaching in American Schools." *The Future of Children* (2007): 1-7.

Heubeck, Tina and Jochen Scheuer, "Incentive Clauses in Players' Contracts in Team Sports- Theory and Practice." *German Working Papers in Law and Economics* 2003 (2003): 1-26.

Hirsch, Barry T., David A. Macpherson, and John V. Winters. "Teacher Salaries, State Collective Bargaining Laws, and Union Coverage," prepared for American Economic Association Meetings, San Diego (January 6, 2013): 1-26.

Jacob, Brian A. "The Effect of Employment Protection on Worker Effort: Evidence from Public Schooling." *National Bureau of Economic Research* Working Paper 19655 (2010): 1-40.

Joyce, Robert P. *The Law of Employment in North Carolina's Public Schools.* Chapel Hill: The University of North Carolina at Chapel Hill, 2000.

Lazear, Edward P. "Performance Pay and Productivity." *American Economic Review*

90, no. 5 (2000): 1346-1361.

Little, Olivia, Laura Goe, and Courtney Bell. "A Practical Guide to Evaluating Teacher Effectiveness." *National Comprehensive Center for Teacher Quality* (2009): 1-30.

Lu, Adrienne. "North Carolina Ends Teacher Tenure." *Stateline*, July 29, 2013. http://www.pewstates.org/projects/stateline/headlines/north-carolina-ends-teacher-tenure-85899493655.

Macartney, Huge. "The Dynamic Effects of Educational Accountability." *National Bureau of Economic Research* Working Paper 19915 (2014): 1-47.

McREL, in collaboration with the North Carolina State Board of Education. *North Carolina Teacher Evaluation Process* (2009): 1-50.

National Education Association. *2011-2012 Average Starting Teacher Salaries by State*. Compiled from the Collective Bargaining/Member Advocacy Database (October 31, 2012).

NCDPI Division of Accountability Services. *The North Carolina Testing Program: 2012-2013*. July, 2012. http://www.ncpublicschools.org/docs/accountability/nctpoverview1213.pdf.

Neal, Derek and Diane W. Schanzenbach. "Left Behind By Design: Proficiency Counts and Test-Based Accountability." *Review of Economics and Statistics* 92 (2010): 263-283.

North Carolina Education Research Data Center. *Summary of Data Files Available*. May 31, 2013. https://childandfamilypolicy.duke.edu/pdfs/projects/NCERDC_Full_Data_Summaries_May13.pdf.

North Carolina General Assembly. *Principal and Teacher Employment Contracts*, G.S.115C-325(c-e)(1).

Over, Paul. "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality." *Quarterly Journal of Economics* 113, no. 1 (1998): 149-85.

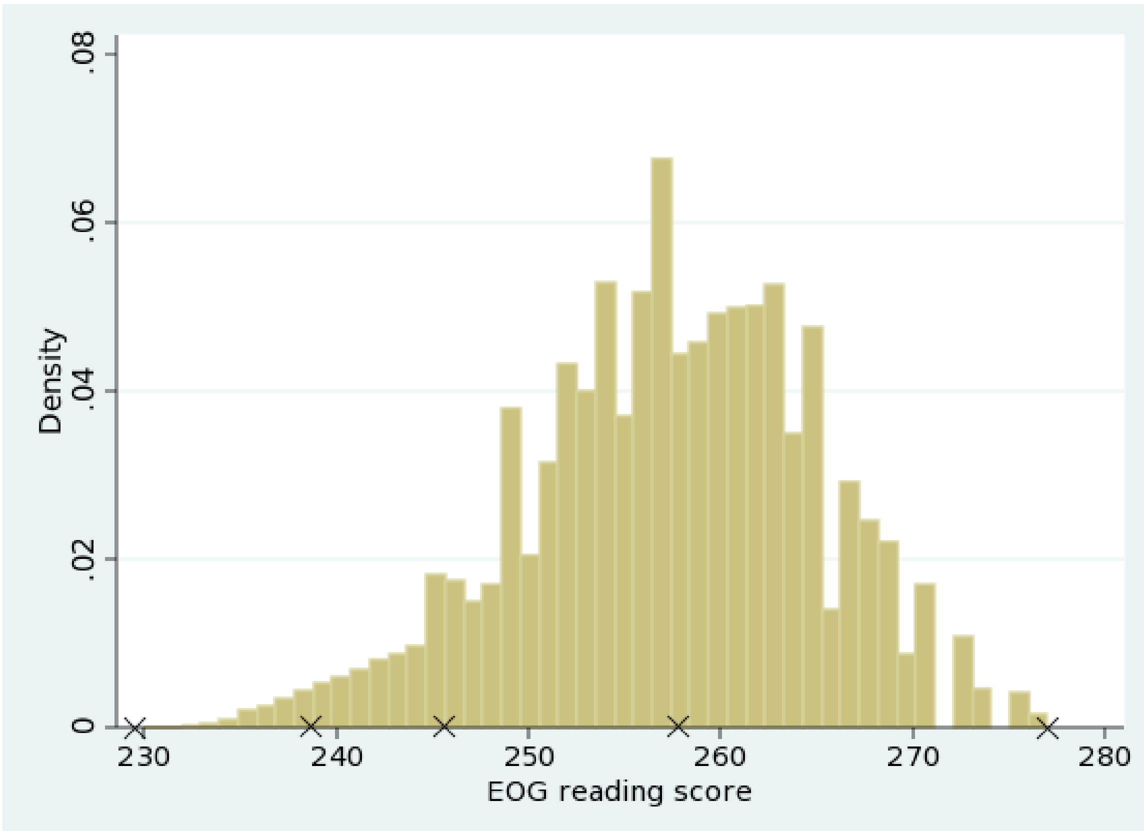
Stephey, M.J. "A Brief History of Tenure." *Time*, November 17, 2008. http://content.time.com/time/nation/article/0,8599,1859505,00.html.

Wiswall, Matthew. "The Dynamics of Teacher Quality." *Journal of Public Economics* 100 (2013): 61-78.

Appendix: EOG Test Score Distributions (Grades 5-8)

Figure A.1

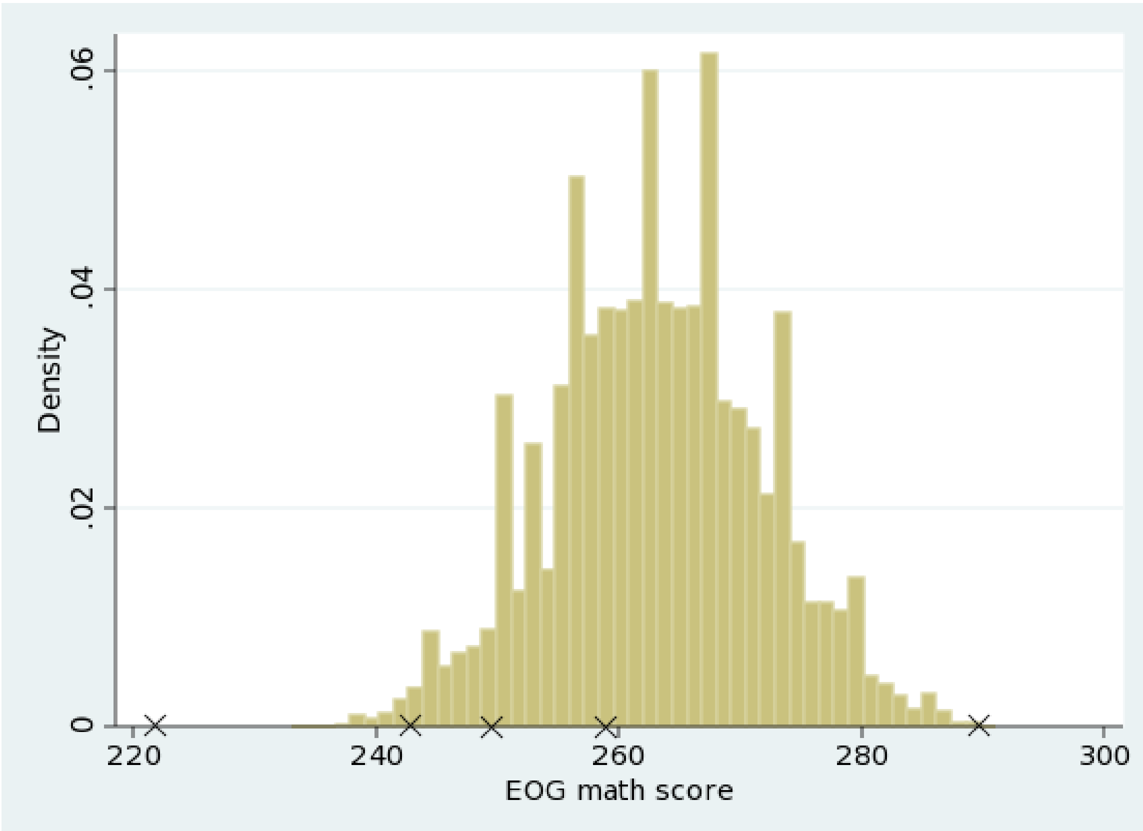
EOG READING SCORE DISTRIBUTION, GRADE 5



SOURCE. – North Carolina Education Research Data Center, EOG Grade 5, 2004

Figure A.2

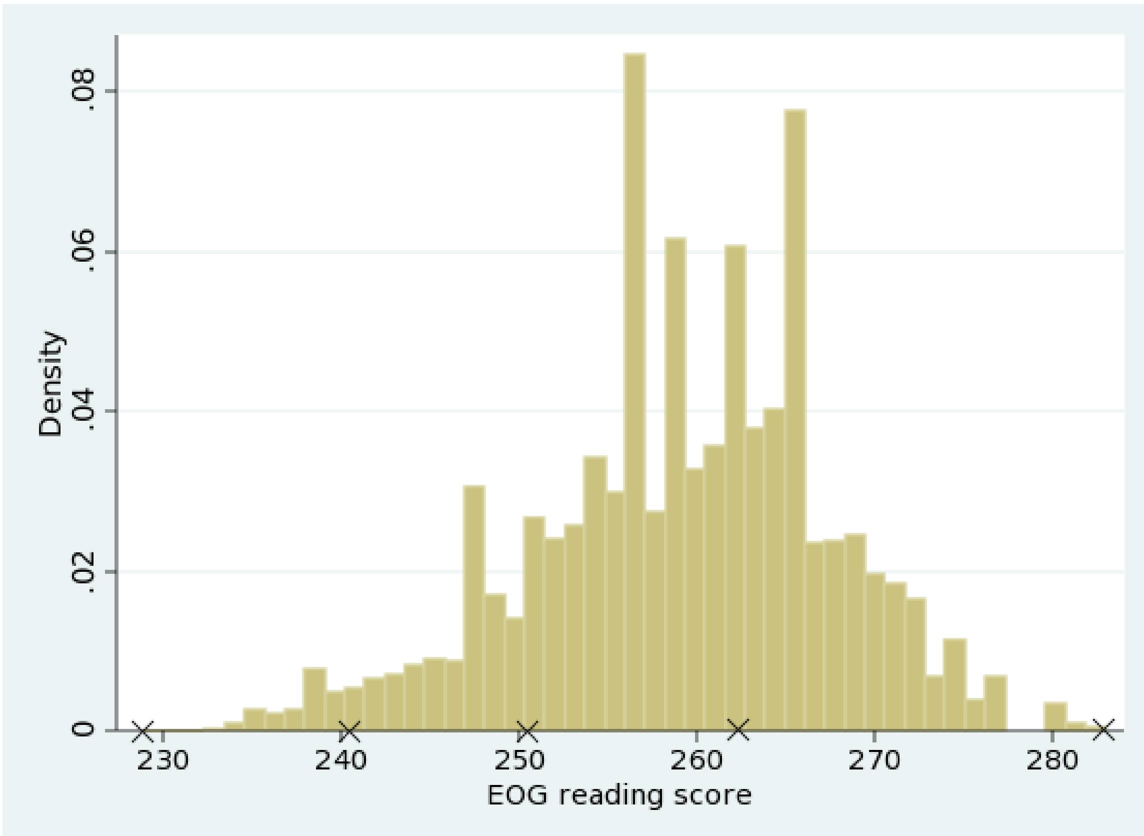
EOG MATHEMATICS SCORE DISTRIBUTION, GRADE 5



SOURCE. – North Carolina Education Research Data Center, EOG Grade 5, 2004

Figure A.3

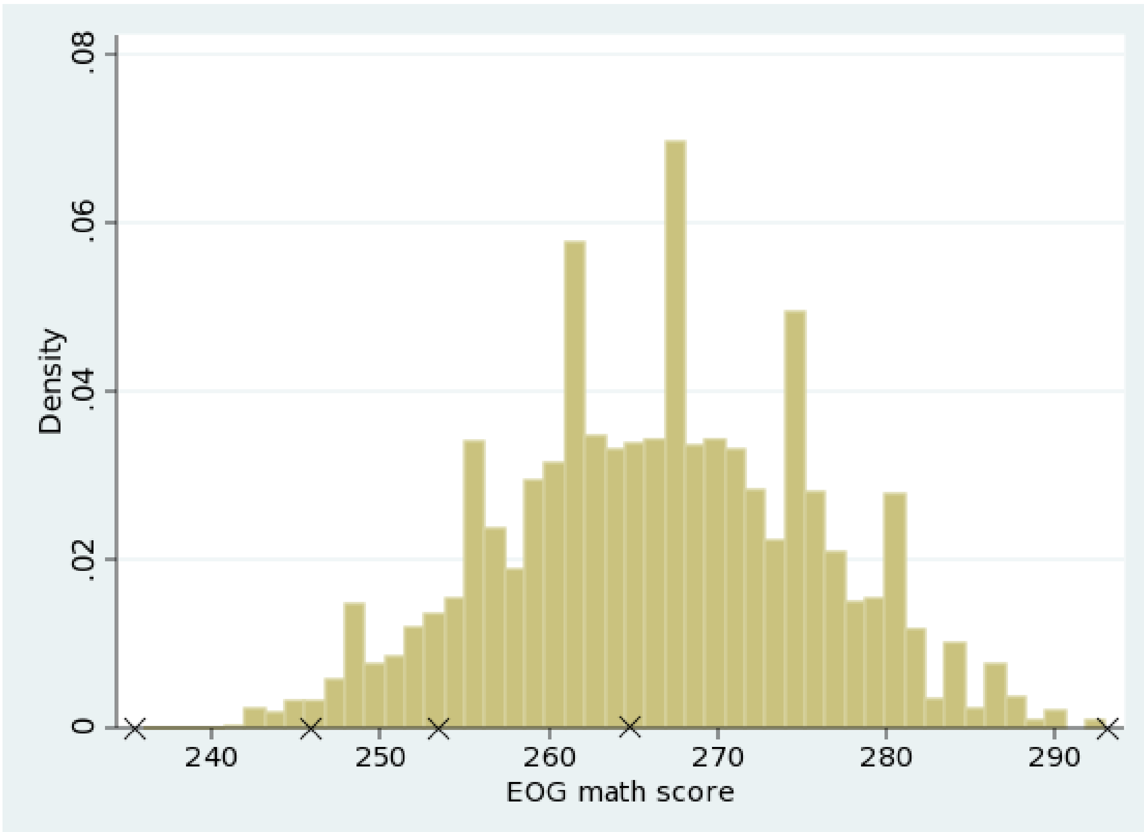
EOG READING SCORE DISTRIBUTION, GRADE 6



SOURCE. – North Carolina Education Research Data Center, EOG Grade 6, 2004

Figure A.4

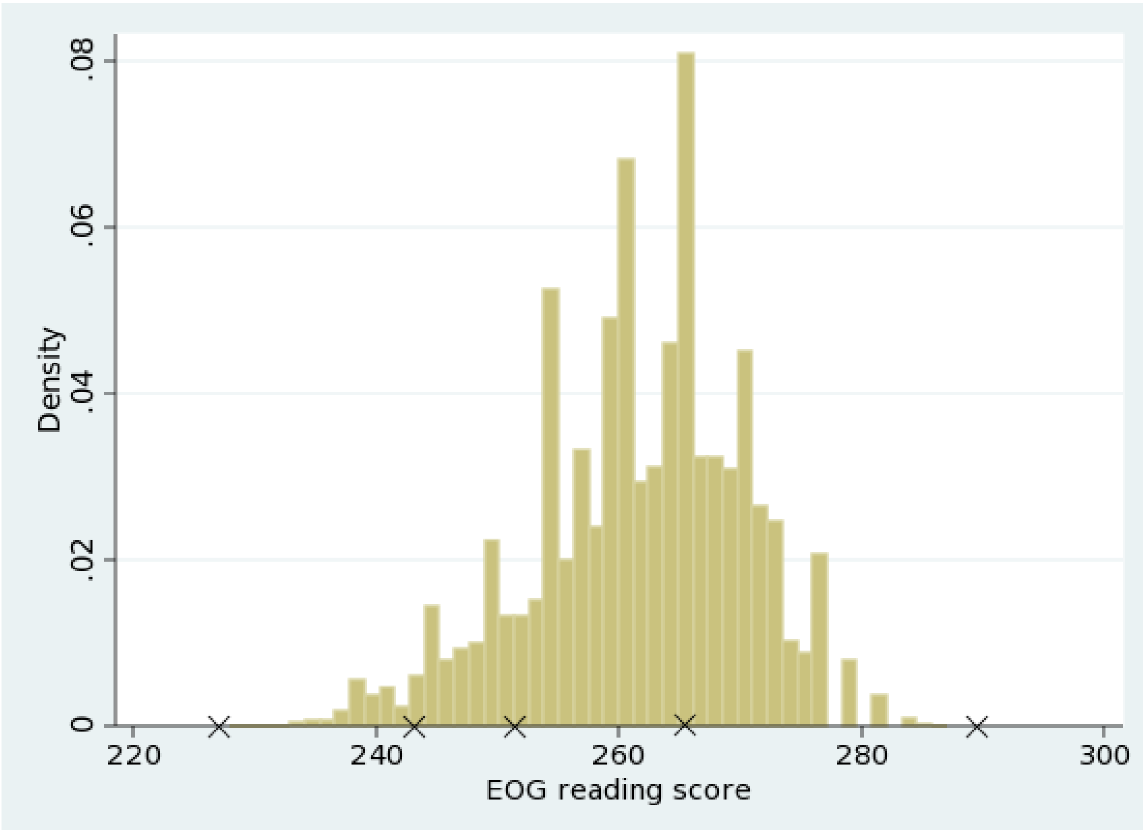
EOG MATHEMATICS SCORE DISTRIBUTION, GRADE 6



SOURCE. – North Carolina Education Research Data Center, EOG Grade 6, 2004

Figure A.5

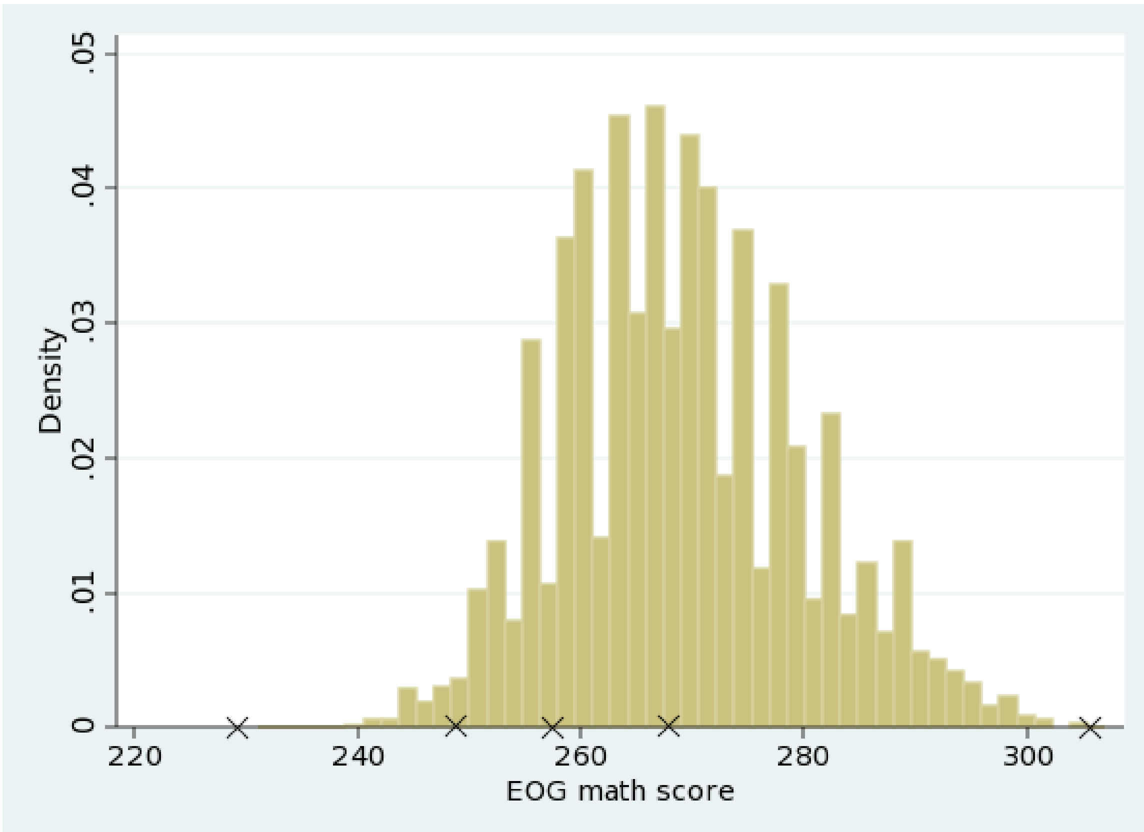
EOG READING SCORE DISTRIBUTION, GRADE 7



SOURCE. – North Carolina Education Research Data Center, EOG Grade 7, 2004

Figure A.6

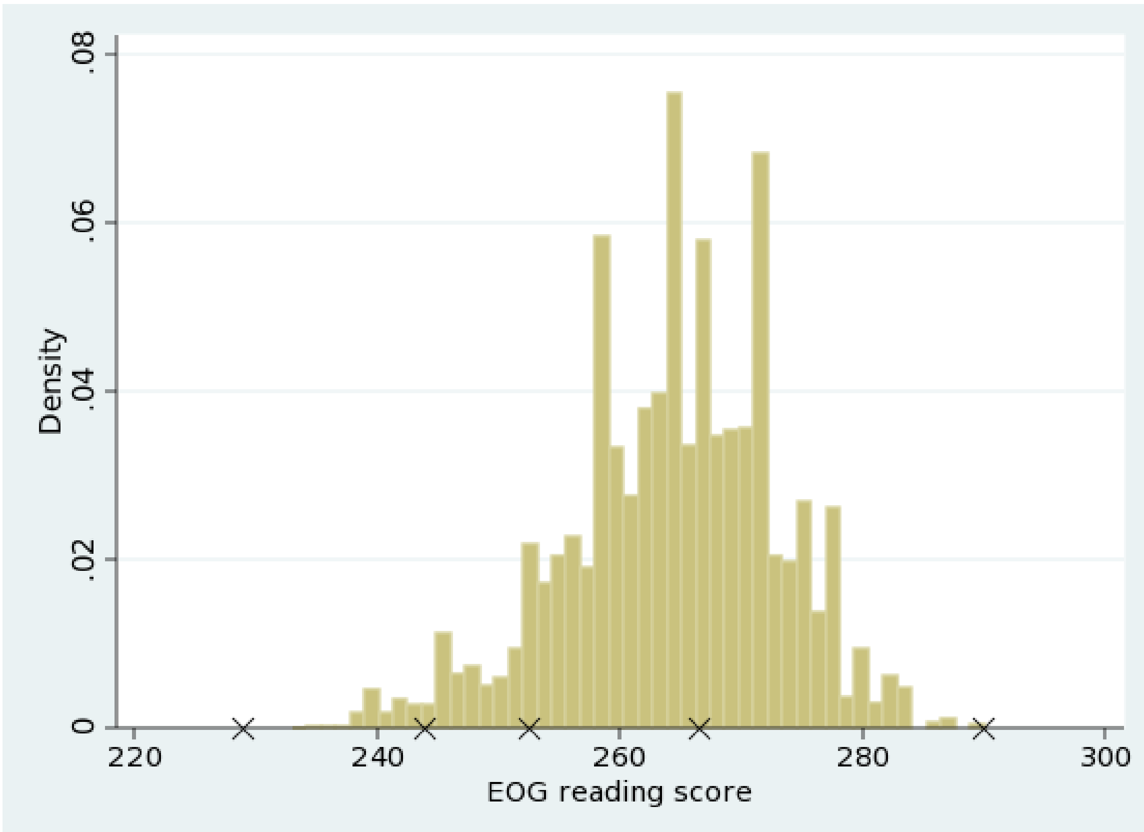
EOG MATHEMATICS SCORE DISTRIBUTION, GRADE 7



SOURCE. – North Carolina Education Research Data Center, EOG Grade 7, 2004

Figure A.7

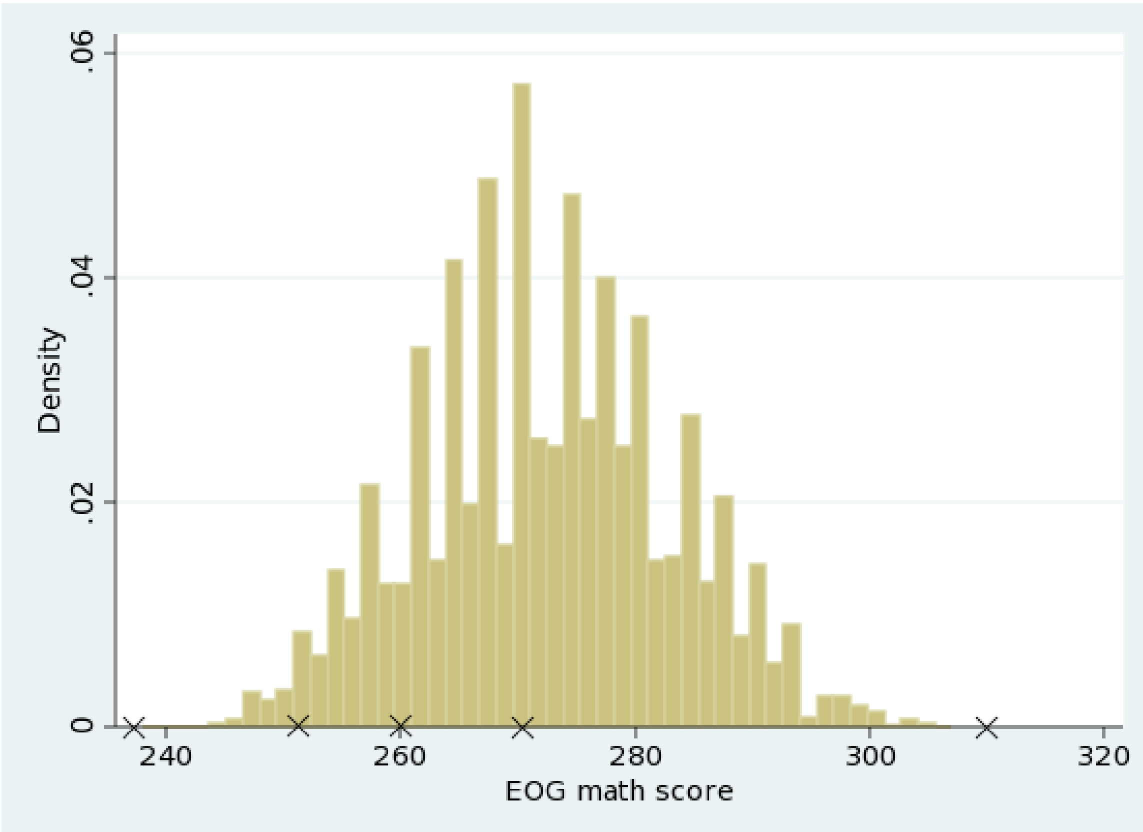
EOG READING SCORE DISTRIBUTION, GRADE 8



SOURCE. – North Carolina Education Research Data Center, EOG Grade 8, 2004

Figure A.8

EOG MATHEMATICS SCORE DISTRIBUTION, GRADE 8



SOURCE. – North Carolina Education Research Data Center, EOG Grade 8, 2004