5-2013

Do Teacher Characteristics Affect Student Achievement? Evidence from a Rural State

James V. Shuls
University of Arkansas, Fayetteville

DO TEACHER CHARACTERISTICS AFFECT STUDENT ACHIEVEMENT? EVIDENCE FROM A RURAL STATE

DO TEACHER CHARACTERISTICS AFFECT STUDENT ACHIEVEMENT? EVIDENCE FROM A RURAL STATE

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Education Policy

By

James V. Shuls
Missouri Southern State University
Bachelor of Science in Education, 2004
Missouri State University
Master of Arts in Education, 2007

May 2013
University of Arkansas

Abstract

Using a unique dataset, this dissertation analyzes the relationship between observable teacher characteristics and teacher effectiveness. Effectiveness is measured as a teacher's ability to improve student achievement on a standardized test. This analysis focuses on teachers of math and science at the elementary and middle school levels, as well as teachers of algebra, geometry, and 11th grade English language arts. The sample is drawn from Arkansas, a state that has a much more rural population than other states and cities where this type of analysis has been conducted in the past. For the analysis I am unable to link teachers directly to students, but I can match teachers and students to a course within a school. Thus, I generate a value added score at the school-course level and attribute that score to each teacher in that school, who teaches that course. I then regress observable teacher characteristics on the school-course value-added measure. In this way, I analyze the relationship between of a host of characteristics and student achievement. The observed characteristics include: certification route, licensure exam scores, experience, graduate degrees, and in some cases college degrees. I conclude with policy recommendations.

This dissertation is approved for recommendation
to the Graduate Council.

Dissertation Director:

_____

Dr. Jay P. Greene

Dissertation Committee:

_____

Dr. Gary W. Ritter

_____

Dr. Patrick Wolf

Dissertation Duplication Release

I hereby authorize the University of Arkansas Libraries to duplicate this dissertation when needed for research and/or scholarship.

Agree _____
James V. Shuls

Refused _____
James V. Shuls

Acknowledgements

The completion of this dissertation and my Ph.D. program in education policy is a tremendous accomplishment. However, I cannot accept all of the credit for this achievement. I would not have made it to this point without the advice, wisdom, support, and guidance of many individuals along the way. I would be remiss if I did not take a moment to thank each of these individuals.

For starters, I must thank my family: Doug Shuls, Elaine Lankford, Belinda Tenny, Michelle Rohan, and John Shuls. Thank you for always believing in me. I love you all.

My best friends, Johnny Suermann and Shaun Simms, thank you for your unwavering support and your unyielding friendship. I could not have asked for better friends.

I must also thank Eric Smith. A year after high school I was working full time and had no intention of going to college. One evening in the summer of 2000, Coach Smith came in to the restaurant where I worked with some of his friends. That evening he shared many pieces of advice, some of which I continue to share with others. He asked me what I planned to do with my life and encouraged me to go to Missouri Southern State University, his alma mater. Within a matter of weeks, I moved across the state and enrolled in college. His prompting set me on the path that I am today and for that I am truly grateful.

Upon graduating from Missouri Southern, I had the opportunity to work with many great teachers in the Republic R-III School District. I am thankful for many of those individuals, but none more than my *partnerooni*, Kelli Torbett. Kelli served as my mentor during student teaching. Later, when I began teaching 5th grade, she and I shared students as teaching partners. I am thankful for her mentoring and support in the classroom, but more importantly for her friendship.

Along the way, there have also been many professors who have made a lasting impact on my life—either through increasing my knowledge or changing the way I think about issues. Dr. John Hail and Dr. Cindy Wilson are two of those professors. While at Missouri State University working on my master's degree, Dr. Hail and Dr. Wilson served on my thesis committee. They continually challenged me and encouraged me to pursue an advanced degree. I am particularly indebted to Dr. Hail, who was one of the most effective an fun professors I have ever had, and who has also become a great friend.

When I left teaching to pursue this degree, I knew it would be an adventure. I am glad that I was able to take that adventure with some remarkable individuals, my classmates: Stuart Buck, Trae Holzman, Dan Bowen, Mike McShane, and Lynn Woodworth. Gentlemen, thank you. In those early days as we studied in the *Nerdery*, I was not sure if I would make it through. I certainly would not have made it, if it were not for all of you. Now as I look back, I would not trade that experience for the world, because it brought us all closer together. I am happy to report that you all are most certainly not a bunch of DBs.

My experience at the University of Arkansas was not just enriched by my cohort, but the entire Department of Education Reform. It was an amazing place to work and to learn. Anna Jacob, thank you for your jokes. While completing my degree, I was fortunate to work for Dr. Robert Maranto. Bob, thank you for being an excellent mentor and friend. I am grateful that you treated me like a valued colleague, not simply as a student. I can honestly say that you taught me many things that I simply could not have learned in the classroom. However, one thing will always remain a mystery. How do you find anything in that office?

In addition to learning from some of the preeminent scholars in education policy, I was fortunate to have them serve on my dissertation committee: Dr. Patrick Wolf, Dr. Gary Ritter,

and Dr. Jay Greene. Dr. Wolf, you are the picture of a scholar. In the classroom you were an effective instructor, always paying attention to the important details. Throughout this dissertation process, your comments were helpful and insightful. I am grateful that I have had the opportunity to learn from you.

I had the opportunity to work as a research assistant for Dr. Ritter and the Office for Education Policy my first year in graduate school. I have now had the opportunity to learn from him and to work with him on scholarly papers. I have appreciated his continual guidance and support. He is particularly skilled at tempering my statements when I get a little too passionate. Dr. Ritter, thank you for commitment to helping me grow as a professional. I am fortunate to have you serve on my dissertation committee and to call you a friend.

I can say undoubtedly that I would not be where I am today if it were not for Dr. Jay Greene. As a master's student I picked up this red and green book at a local book store. It was called "Education Myths." That book challenged the way I looked at issues and made me interested in education policy research. I soon became aware that Jay was the head of the Department of Education Reform and would soon be enrolling candidates in a brand new doctoral program. I jumped at the opportunity to learn from Dr. Greene. Often times when you read an author, you put that person up on a pedestal. Invariably, these individuals never live up to your expectations when you meet them in person. This was not the case with Dr. Greene. I can honestly say that after having him as a professor, a mentor, and a dissertation chair that I respect him more. While I will always hold him in the highest regard, I am thankful that he taught me that he is just an old guy with a mustard stain on his tie. Thank you Dr. Greene, for helping me transition from a consumer of research to a producer of research.

Finally, I must thank my beautiful wife Crissy Shuls for all of her support. When we met during our freshman year at Missouri Southern she probably never imagined I would still be in college 13 years later. Thank you for standing with me as I pursued this dream. There were many hard times along the way, but together we always managed to make it through. This degree is as much your accomplishment as it is mine; because without you, I would not be here. I love you.

Last, but not least, I must thank my children: Tryston Scott, Taryn Mackenzie, and Truman Michael. This dissertation is dedicated to you. Thank you for bringing so much joy to my life. Thank you for being such amazing children. When I look at you, I cannot help but marvel at the wonderful creation God has made. I see so much promise in each of you. I know not what path the Lord has for you, but I am honored that he entrusted me to start you down that path. Let this accomplishment be a testimony to you; if someone as silly as your father can achieve this, then the possibilities for you are endless. Whatever life has in store for you, always remember to *Be good and Do Good.*

James V. Shuls

April 2013

Table of Contents

Chapter 1 - Introduction

Since the beginning of organized schooling, parents have sought to get their children into a good teacher's classroom. Before studies were conducted and regression analyses were run, parents knew intuitively that some teachers were better than others. Over the past few decades, scholarly research has begun to verify what parents have known all along; that teacher quality matters. President Obama reiterated these findings in a 2009 town hall meeting when he remarked that the "single most important factor in the classroom is the quality of the person standing at the front of the classroom."

For obvious reasons, improving teacher quality is a worthwhile pursuit; one that policymakers and academics have pursued for decades. Though there are many ways to do this, most proposed ideas for improving the quality of teachers fall into one of two camps: quality control through entrance or through deselection. The first camp seeks to limit entry into the profession through some form of licensing, often requiring teachers to pass a test of skills or to complete specified coursework. The idea here is that by training and equipping teachers on the front end, only the most qualified and effective individuals will enter the classroom. The most prominent member of this camp is Linda Darling-Hammond (1997), who wrote "[T]eachers who are fully prepared and certified in both their discipline and in education are more highly rated and are more successful with students than are teachers without preparation, and those with greater training in learning, child development, teaching methods, and curriculum are found to be more effective than those with less" (p. 10). Notable proponents in this camp are Dianne Ravitch (2010) and Marilyn Cochran-Smith (2010) among others.

The second camp suggests that entrance to the profession should be less restrictive, with all types of individuals entering the classroom. By allowing easy entrance into the profession, the

applicant pool for teaching positions is expanded. That means administrators can choose from a larger pool, possibly leading to the selection of higher quality teachers. This camp also believes that low-performing teachers should be weeded-out, or deselected from their teaching post. The most notable scholar in this camp is Eric Hanushek (2010), who has documented that replacing low-performing teachers with teachers of average quality can have a large impact on student achievement. Dale Ballou and Mike Podgursky (2000) and Kate Walsh (2001) are other leading scholars in this camp. In the United States, policymakers have tended to favor the former position over the later.

One way to assess these competing hypotheses is to examine a fundamental assumption of the certification argument. The claim that certification enhances quality is hinged on the belief that the quality "screens" actually reflect quality.. That is, the tests and training required to obtain entrance into the teaching profession must be accurate indicators of effectiveness if they are to improve the quality of the workforce. Thus, the purpose of this study is to assess whether the licensure screens used for the certification of teachers in Arkansas are predictors of teacher quality in regard to effectiveness at improving student achievement scores. Additionally, I will examine other observable teacher characteristics and their impact on student achievement. The primary focus of this study, however, will be to examine the differences in terms of performance between traditionally and alternatively certified teachers.

Although there have been a number of studies on the impact of licensure route and other observable teacher characteristics on student achievement, these studies have been conducted in few states and the results are not yet conclusive, especially in regards to certification. Furthermore, most rigorous studies have been conducted in locations that do not reflect the makeup of the teacher workforce in other states, especially Arkansas. That is, Arkansas is much

more of a rural state than New York, Chicago, Houston, or even North Carolina; all locations of previous studies.

I expand on the body of knowledge by providing analyses in an environment that is much different than New York City or North Carolina. Very few prospective Arkansas teachers attended a highly competitive university. No Arkansas college or university is listed in the highest competitive rating category by *Barron's Guide to the Most Competitive Colleges*. Arkansas' flagship, land grant institution, the University of Arkansas, is listed as number 134 on *U.S. News and World Report's* National University Rankings (2013). This ranking takes into account the overall quality of the university, not necessarily the college of education. No other Arkansas institution of higher learning is listed on the national university rankings. Arkansas also has a smaller percentage of the adult population with college degrees, 18.9 percent compared to 31.8 and 25.8 percent in New York and North Carolina, respectively (US Census). Thus, Arkansas has a smaller pool of individuals to which alternative teacher certification is even viable. In short, the findings from Arkansas may generalize to parts of the United States, particularly rural states, in a way that New York and North Carolina cannot.

First, it is important to understand the rationale for licensing and a bit about the historical development of teacher certification in the United States. Therefore, in this section I will provide a brief introduction to the theory of certification in general and the historical development of teacher certification practices in the United States. Next, I will describe the pathways to licensure in Arkansas. I close this chapter by outlining the research questions I will explore in this dissertation.

Certification and Occupational Licensing

Throughout history, governments have sought to regulate some occupations by enacting laws that limit entry into the profession. For example, the medical profession has been regulated in some way since at least the 13th century (Daly 1998). A common method of regulating a profession is to require occupational licensing of individuals that wish to carry out the trade. Kleiner (2000) defines occupational licensing "as a process where entry into an occupation requires the permission of the government, and the state requires some demonstration of a minimum degree of competency" (p. 191). This type of occupational licensing has existed in America since colonial times, but the prevalence of licensing exploded in the later part of the 1800's (Rottenberg 1962). Today, roughly 29% of the entire United States' workforce is comprised of workers required to have a license (Kleiner & Krueger 2010).

A distinction is typically made between occupational licensing and certification. Kleiner (2000) suggests, both licensing and certification require demonstration of basic skills, but consumers can choose whether to hire a certified person or not, whereas it is illegal for someone without an appropriate license to perform the regulated job. In education the two terms are often used synonymously, as they are in this dissertation; but based on the descriptions above, teacher certification falls more in line with traditional occupational licensing. In 2010, roughly 90 percent of all students were in traditional public schools (National Center for Education Statistics 2010a and 2010b). Most public schools will not hire a teacher unless the teacher has a valid teaching license or certificate. Many students in the United States do not have educational options outside of public schools; that means they are unable to choose between a certified and uncertified teacher. In essence, individuals are required, in most circumstances, to send their child to a school with an appropriately licensed teacher.

In 2001, The Abell Foundation noted that every state had regulated the teaching profession with certification or licensure requirements (Walsh 2001). This was prior to the passage of No Child Left Behind (NCLB), which required schools to hire *highly qualified* teachers. Understandably, many states interpreted the *highly qualified* provision as synonymous with certified. Thus, traditional public schools seek certified applicants for vacant positions and are typically not permitted to hire uncertified teachers unless there are no qualified candidates with certification or a waiver is granted.

Although there are alternative ways of entering the classroom that do not follow the traditional path, nearly every route to teaching requires the applicant to obtain a license at some point. Even alternative pathways to the classroom require the prospective teacher to obtain some form of licensure before he or she is allowed to be a teacher of record in a classroom. Thus, even when districts hire alternatively certified teachers they are still hiring individuals with a state issued certification. It is only in the direst of circumstances when a school district will hire an individual without the appropriate credentials and even then, the teacher typically must begin working towards certification. For all of the aforementioned reasons, teacher certification falls more under the definition of occupational licensing.

The rationale for licensing is often based on consumer protection. Policymakers want to keep charlatans out of the profession and ensure a minimum quality of goods or services for consumers (in this instance, students and families). In theory, ensuring minimal quality helps consumers by protecting them from potential hazards and by improving the average quality of goods or services received. This can only happen, however, when the screens for licensure are appropriate measures of effectiveness. An appropriate screen would allow into the profession only those individuals who have the skills and expertise to do the job well. An ineffective screen

would open the certification up to both type one and type two errors. That is, it would falsely keep some highly-effective individuals out of the classroom, while simultaneously letting some individuals into the profession who have no business teaching.

Although consumer protection and improving quality are often cited as the primary reasons for licensing, it must be noted that the practitioners of the regulated trade also benefit from increased regulation. It is a simple, widely accepted theory in economics that when the supply of a good or service is limited, the prices rise. Limiting entry to the profession is exactly what licensing does; and as Friedman and Kuznets (1945) demonstrate, increased licensing allows regulated professionals to charge higher fees for their services. An increase in wages occurs because the supply of workers in the regulated field is decreased. Limiting the supply of providers decreases the supply of available labor, thus increasing the price, or wage paid, for the professional already practicing the profession (Kleiner 2000; Kleiner & Krueger 2010; Rottenberg 1962, 1980).

Adam Smith identified this phenomenon of increasing prices through regulation centuries ago. At that time, it was common in many townships to require tradesmen to undertake a lengthy apprenticeship before practicing a profession independently. The lengthy apprenticeship discouraged some from entering these professions. What's more, many of these professions limited the number of apprentices a master could have. As Smith (1863) states, "the intention of both regulations is to restrain the competition to a much smaller number than might otherwise be disposed to enter the trade" (p. 54). These rules were ultimately codified into laws in many places, with the practitioners of the trade actively lobbying for passage of the very regulations that benefit their business.

Though formal apprenticeships are no longer the norm, occupational licensing still limits access to a profession in a number of ways. Three of the most common measures used to restrict entrance to a profession are:

- Entrance exams
- Education credentials
- Residency requirements

Entrance exams are possibly the most common requirement for many occupational licenses, in the legal field, prospective lawyers are required to take the Bar exam; in teaching most states require a test of basic knowledge and content examinations. Examinations are useful in limiting entry to the profession, because the passage marks do not have to be static. When the number of individuals in a field grows too high or drops too low, the score required to pass can be adjusted to increase or decrease the number of entrants to the profession (Kleiner 2000).

Requiring individuals to obtain specific education credentials is another oft used requirement for occupational licenses. Again, the end result is that these requirements restrict entry to a profession (Kleiner 2000). All teachers in every state are required to have a bachelor's degree; while other professions often require a degree in the field in which the individual will be working. As part of the educational experience, individuals are sometimes required to complete some form of residency. The most widely known residency requirements are in the field of medicine, but many occupations require some form of residency. For example, though it may not be thought of as such, student teaching is a form of residency. Before a teacher can graduate from many education programs and before he or she may obtain a standard teaching license he or she must complete a required student teaching experience, typically 12 weeks in duration.

Each of these -- examinations, education credentials, and residency requirements -- act as a barrier to entry of a profession. A simple definition of a barrier to entry is any obstacle that does not allow an individual to enter that market freely. For obvious reasons, these three requirements are barriers to entry. More than that, these barriers also increase the costs of entering a profession (Rottenberg 1962). It costs the individual money when he or she must pay the cost of tuition and exam fees in order to meet licensing requirements. These requirements also require individuals to sacrifice income that could be earned while fulfilling the requirements. That is, in the four years it takes an individual to earn a degree, they could be working; the same could be said for the time they spend studying or preparing for the licensing exams.

Given the impact on potential wages, practitioners have significant motivation to minimize competition in their field by limiting the number of individuals entering their profession. Indeed, evidence suggests occupational licensing increases wages approximately 15%, roughly the same impact as unionization (Kleiner & Krueger 2010). Thus, as Rottenberg (1962) suggests, calls for licensing "are almost invariably made by practitioners of those trades, not by consumers of their services; and standards are established and examinations conducted by boards of examiners composed of practitioners" (p.4). For these reasons Gellhorn (1976) believes, "That restricting access is the real purpose, and not merely a side effect, of many if not most successful campaigns to institute licensing schemes can scarcely be doubted" (p. 11).

Brief History of Teacher Licensing

Each day millions of parents entrust their children to the care of teachers in public schools. Parents want to know that the individual standing in front of the classroom is competent and capable. To the general public the rationale for teacher licensing seems almost self-evident.

Individuals that desire to teach should have some requisite knowledge and ability to actually educate children. As Goldhaber (2011) writes, "those who wish to become public school teachers ought to have a basic skill set related to teaching" (p. 316). To fully understand the teacher certification laws of today, it helps to have some context about how these policies have developed throughout history.

The concern for quality teachers is nothing new. Indeed, throughout history people have been concerned about the education of their children. The earliest record of teacher credentialing dates to the Roman Empire (Angus 2001). The Romans, like nearly every culture since, believed that teacher quality mattered. Colonial Americans were just as interested in teacher quality, although typically for religious reasons. After all, much of a child's education at the time was religious in nature. Take, for example, The Old Deluder Satan Law, enacted by Massachusetts in 1647. The law stated the "chief project of that old deluder, Satan, [is] to keep men from the knowledge of the Scriptures" (O'Donnell 2010, p. 676). Therefore, it was a necessity to teach children how to read so they could read the Holy Scriptures. Accordingly, the law required every town comprised of more than fifty households appoint a teacher to instruct children to read and write.

The most important trait of a potential teacher in this era was that of character. For this reason, "In colonial America, it was common for communities to require that anyone proposing to teach be approved by one or more of the local ministers" (Angus 2001, p. 13). Upon the minister's approval, the townspeople often credentialed teachers. This was a local endeavor for several reasons. First and foremost, the authority to educate children came from local governments. Most colonies had yet to establish or fund public schools. Thus, each community could decide for itself to have a school and what the education of that school would entail.

In those days, students typically did not attend school for more than the elementary grades. Thus, possessing a degree beyond high school was typically not thought to be necessary. Moreover, teacher training in the early days of America was not as highly specialized as it is today. Though some training programs existed as early as the 1780s, training typically occurred in teacher academies and seminaries. These training programs were privately run, and often privately funded. However, in some instances these institutions were publicly funded; like they were in New York during the 1820s and 30s (Angus 2001). In rural areas where access to higher education was limited, the training of teachers often took place in public high schools (Angus 2001).

Common Schools and Common Regulations

Throughout the 1800's, state governments increasingly assumed licensing authority. As early as 1825, important political figures began calling for an improved teacher labor force to meet the needs created by a democratic representative government (Harper 1839). These calls would eventually lead Massachusetts to establish a state board of education in 1837. Horace Mann was appointed as the first secretary of the board, a role he assumed for the next eleven years (Reese 2011). Discouraged by the state of free public education in the United States, Mann and other education reformers "called for longer school terms, better daily attendance, school consolidation, the professional training of teachers, and a host of other improvements" (Reese 2011, p. 12). These other improvements included a statewide system of free public schools supported by tax dollars. These free public schools became known as "Common Schools."

The Common School movement changed public education in the United States and hastened the transition of licensing authority from local jurisdiction to the state (Angus 2001). In many instances, education went from being a local prerogative to a state mandate. This resulted

in the number of students enrolled in public free schools increasing dramatically. Also increasing was the number of individuals completing high school. With the increase in students, more individuals were needed to fill the role of teacher in the burgeoning systems of public schools. As a result, teacher training programs sprouted throughout much of the Northeast and Midwest.

"On July 3, 1839, the first state-supported school for the exclusive purpose of preparing teachers was opened at Lexington, Massachusetts" (Harper 1939, p. 9). This was the first Normal School (Angus 2001). As the demand for teachers trained to deliver the common education offered in common schools grew, teachers were increasingly trained in normal schools. The first western normal school was opened in Michigan in 1853 and Illinois quickly followed suit (Harper 1939). In fact, the normal school in Illinois marked a historical change in the history of teacher preparation. Having no state college, the normal school was dubbed *The Illinois State Normal* University, a college for teachers. Not to be outdone, Pennsylvania enacted a law in 1857 that created normal schools in twelve separate districts throughout the state (Harper 1939). "Between 1860 and 1900, the idea of training teachers in specialized, tax-supported institutions spread fairly rapidly throughout New England and into the Midwest. By the close of the nineteenth century, there were 127 state-supported normal schools and a slightly larger number of private normal schools" (Angus 2001, p. 5-6). Many of these normal schools developed into teacher's colleges and subsequently into many of the colleges and universities of today.

Existing colleges and universities also jumped into the teacher training business. The University of Iowa was the first to establish a department of education within the university in 1855 (Angus 2001) and the first permanent chair of pedagogy in 1873 (Harper 1939). Prior to this, professors of education at some universities acted as one man departments of education. The education department at the University of Iowa would go on to be the largest department in the

university's system, enrolling half of all students from 1858 to 1864 (Angus 2001). By 1890, roughly one fourth of all colleges and universities prepared teachers (Angus 2001).

With the growth of teacher preparation programs and the increased demand for teachers came calls for ensuring a minimum level of teacher quality. At first, locals decided the content of teacher examinations. As centralization of education increased throughout the 1800's, so did the content of teacher examinations. Tests that had once been primarily oral examinations transitioned to written exams (Angus 2001). The subject matter also changed from character driven to content driven questions.

In 1834, localities in Pennsylvania began requiring teachers to pass a test of basic skills before they could be credentialed as a teacher (Ravitch 2003). In 1841, New York began requiring teachers to be tested on a variety of basic skills, namely: reading, writing, arithmetic, geography, and history (Angus 2001). Not long after, in 1843, New York was the first state to institute statewide certification for teachers (Angus 2001). Roughly ten years later, Indiana and Pennsylvania followed suit. These exams were meant to ensure quality.

Initially, professional educators sought to limit the use of testing to license teachers, because it was seen as a way to provide a back door into the classroom. Instead, professional educators preferred formalized teacher education training programs as a prerequisite for licensure. The notion that specialized knowledge is needed to teach spilled over into the teacher testing business. Pennsylvania, an early adopter of teacher licensure exams, became the first state to include "professional knowledge," a practice that continues today in most states (Angus 2001).

The history of education in the United States is one of increasing government regulation and the "professionalization" of teaching. Throughout the 19th century, much of the education reform agenda was fueled by the Common School movement. This was, in part, an effort to

provide a system of free public education to the masses of immigrants flooding the country. Education began as a local endeavor, but quickly became a bureaucratically controlled industry, with the influence of education professionals. Teaching became a specialized skill, acquired through formal training. Teachers were increasingly trained in specialized teacher prep programs, beginning with Normal Schools, teacher's colleges, and then departments of education in universities. By 1953, the number of states requiring a four year degree for elementary school teachers was 25, while 45 states required four or more years of college for high school teachers (Angus 2001).

Throughout much of the 1900s the pathway to the classroom became increasingly narrow and by the early 1980's prospective teachers in the U.S. were typically required to graduate with degrees in education if they wanted to teach in a K-12 public school. In 1983, teacher licensure began a new course, with the advent of alternative certification. New Jersey led the nation by establishing the first alternative certification program in 1983 (Walsh & Jacobs 2007). A variety of factors, including concerns about teacher quality and teacher shortages, led to the creation of alternative certification routes, but the creation of alternative pathways to the classroom was met with strong resistance by universities and teachers' unions. Despite opposition from the education establishment, COMMA many states began implementing alternative pathways as a means to reduce teacher shortages.

Today, nearly every state has an alternative route to teacher licensure; however, many of these paths are not dissimilar to traditional pathways. Walsh and Jacobs (2007) note most alternative programs are operated by colleges of education. The programs often have the same requirements of student teaching, specified coursework, and pedagogical knowledge found in traditional licensure programs.

Teacher Certification in Arkansas

Although there are many commonalities among certification programs across states, each state has a somewhat unique system of teacher licensure. Since this study examines the impact of teacher certification on student achievement in Arkansas, it is important to note the specifics of Arkansas' pathways to teaching. In Arkansas, there are a number of different pathways to the classroom. Though the distinctions can be a bit blurry at times, for the purposes of this study, each of these pathways will be sorted into traditional and alternative categories. The distinctive difference between traditional and alternative pathways in Arkansas is student teaching.

The Traditional Route

The key element to being traditionally certified in Arkansas is completing a student teaching experience. In the most common traditional approach, teachers earn a bachelor's degree in an approved education program. This could be a bachelor's in education from a college of education or it could be from a degree in a content area, working in conjunction with the college of education. Individuals wishing to enter into one of these degree programs are required to pass a test of basic skills, the Praxis I, typically in their sophomore year. This test covers mathematics, reading, and writing. Upon passing the Praxis I and finishing required coursework, the prospective teacher must student teach for a minimum of 12 weeks or 360 clock hours (Kimbrell 2010).

In addition to completing all degree requirements and student teaching, prospective teachers must also pass Praxis II examinations. These examinations differ by content area and grade level, but typically consist of a professional knowledge and a content knowledge exam.

Though the bachelor's route is the most common traditional pathway, it is not the only traditional pathway. Some colleges and universities offer a master's degree with a student teaching component. Students at the University of Arkansas at Fayetteville, for instance, can earn a bachelor's degree in elementary education or for secondary teachers in a content area. Once completing their bachelor's degree, they enroll in a master of arts in teaching program (MAT). The MAT program allows them to complete a master's degree and spend a considerable amount of time student teaching. Since teachers in this route take courses in a college of education and complete a student teaching experience, I include MAT students as traditionally trained teachers.

Other teachers I consider to have a traditional licensure include individuals who have a certification in another state and obtain licensure in Arkansas via reciprocity. It very well could be the case that some of these individuals obtained an alternative license in their previous state, but unfortunately my data do not allow me to obtain that information. Therefore, since these individuals have some experience before entering an Arkansas school, I consider them traditionally certified.

Alternative Certification

Arkansas first began implementing an alternative pathway to the classroom in 1988 (Feistritzer & Chester 2000). Today, the alternative certification route consists of individuals who enter the profession via the Non-Traditional Licensure Program (NTLP) or Teach for America (TFA). The NTLP is a way for individuals with a bachelor's degree outside of the field of education to enter the field. To do so, they must have a grade point average of at least 2.5 or 2.75 in their last 60 hours. Additionally, they must pass the Praxis I test of basic skills and for some licenses they must complete some minimal coursework. In addition to these requirements, individuals must be employed by a school district by September first to be admitted to the NTLP.

This is an important distinction, because one cannot earn an alternative license without being employed. Once in the NTLP, the pre-service teacher is awarded a one year provisional license, which can be renewed up to three years. Within this time, the prospective teacher must complete summer instructional modules, be mentored for two hours every two weeks, develop a professional portfolio, and pass the Praxis II licensure exams. Upon completion of all of the requirements, the teacher can earn a standard teaching license.

An additional pathway to certification is provided for members of Teach for America. Teach for America (TFA) is a national organization, which recruits high-performing individuals to work in hard to staff schools. Most TFA teachers in Arkansas are in the Mississippi Delta region of the state. Once accepted, TFA conducts a six week training session for their members and then places them in the classroom. However, to be awarded a two year provisional license, TFA members have to pass the Praxis I and Praxis II tests. Until recently, TFA members were required to go through other pathways, such as the NTLP or a master's program, to earn a standard license. Now, TFA members have a more direct pathway to licensure because TFA's training and mentorship have been recognized as a licensure route by the state.

Research questions

Arkansas' pathway to the classroom is fairly restrictive. Those who do not earn a bachelor's degree from an approved teacher training program and pass all licensure exams must enter the field through the NTLP, TFA, or by earning a master's degree. In any event, they must pass all licensure exams regardless of the pathway they take to the classroom. Certification status and performance on licensure exams should be indicators of effectiveness, since they are requirement for those wishing to teach. Since licensing restricts the pool of potential teachers, it is important to evaluate whether the licensure screens are adequate measures of effectiveness.

This dissertation will examine the following research questions regarding licensure screens, licensure routes, and observable teacher characteristics.

(1). Do traditionally certified teachers generate significantly larger increases in student achievement on standardized exams than alternatively certified teachers?

a. Do they generate the significantly larger gains on the Benchmark mathematics exam?

b. Do they generate significantly larger gains on the Benchmark language arts exam?

c. Do they generate significantly larger gains on the geometry end-of-course exam?

d. Do they generate significantly larger gains on the algebra end-of-course exam?

e. Do they generate significantly larger gains on the 11^{th} grade literacy exam?

(2). Are teacher licensure exams scores related to gains in student achievement?

a. Are teacher scores on the Praxis I exams related to student achievement?

b. Are teacher scores on the Praxis II professional knowledge exams related to student achievement?

c. Are teacher scores on the Praxis II content knowledge exams related to student achievement?

(3). Do factors that determine teacher salaries impact student achievement?

a. Do teachers improve with experience?

b. Are teachers with a master's degree more effective at raising student achievement than teachers without such a degree?

(4). Do secondary teachers with a degree in the field they are teaching generate significantly larger learning gains than teachers without a degree in the content field?

To test whether teacher certification in Arkansas is beneficial, I would ideally examine the performance of traditionally certified and non-certified teachers. That is impossible, however, since all individuals entering the classroom are required to have some form of credential. This does allow me to evaluate the differences between traditionally and alternatively certified teachers, essentially testing the impact of student teaching and education coursework. Of course the individuals who enter the classroom via the traditional or non-traditional route are not randomly selected. The systematic differences of the individuals entering via each path can be thought of as a product of the certification route. Thus, the comparison of each route is not just an evaluation of the preparation a teacher receives, but an evaluation of the route as a whole. That means the comparison is evaluating the effectiveness of a specific preparation pathways ability to train teachers, as well as the ability to attract talented individuals to the field.

The second set of research questions allows me to explore the effectiveness of Arkansas' licensure screens a bit more. Once again, I cannot compare those individuals who passed the licensure exams with those who did not, because the ones that did not pass were denied entry to the profession. However, I can examine whether individuals who score higher on licensure exams outperform individuals who score lower. If differences are found, it can be presumed that these differences would carry on to the left tail of the distribution.

Unlike the first two research questions, the third question does not deal with certification. Rather, it allows me to explore another important aspect of the teaching profession, compensation. Currently, most teachers are paid by a single salary schedule. Most schedules reward teachers with salary increases for each additional year of experience and for earning an

advanced degree. Since experience and degrees are used for compensation, it implies that they are measures of quality. These research questions allow me to explore the relationship between these traditional measures of quality and one of the key measures of teacher effectiveness -- raising student achievement.

In the fourth research question, I explore the relationship between a teacher's degree and their performance in the classroom. For this analysis, I will see if individuals with a math degree are more effective at raising student achievement in algebra and geometry than are individuals without a content degree. Similarly, I will examine the relationship between having a content degree in English language arts and effectiveness as a high school English teacher.

Paper Organization

This dissertation is divided into five chapters. In the next chapter, I provide an examination of the literature as it relates to impact of observable teacher characteristics on student achievement. In Chapter 3, I provide a description of the data and the research methodology used to answer my research questions. The results of my analyses are presented in Chapter 4. I conclude in Chapter 5 with a summary of the findings and policy suggestions.

Chapter II – Literature Review

In a 2009 town hall meeting President Barack Obama commented, the "single most important factor in the classroom is the quality of the person standing at the front of the classroom." Intuition and empirical research affirm this claim. Indeed, a broad research base has documented the impact a teacher can have on student achievement (Rivkin, Hanushek, and Kain, 2005; Hanushek, Kain, O'Brien, and Rivkin, 2005, Nye, Konstantopoulos, & Hedges, 2004). In the past, teacher quality was judged by observations or by a teacher's credentials. In recent years, however, teacher quality has increasingly been measured in terms of a teacher's ability to increase student achievement. This is often called a teacher's value-added, indicating the impact a specific teacher has in terms of increasing a child's performance on a standardized test.

Until the relatively recent creation of longitudinal datasets, it has been nearly impossible to identify a teacher's unique contribution to student learning. Now, states across the country are developing longitudinal datasets whereby teacher characteristics can be linked to student performance. This has allowed researchers to examine the impact of various teacher characteristics on student achievement in an attempt to identify what makes a teacher effective.

In this chapter, I review the literature regarding teacher impacts on student achievement. The literature discussed is limited to studies that evaluate the impact of observable teacher characteristics on student achievement. These characteristics include: undergraduate degree, master's degree, selectivity of university, licensure exam scores, certification subject, certification route, National Board Certification, and experience. Value-added estimates are sensitive to model specification, and there is considerable debate about the appropriate method for calculating a teacher's value-added. In this review, I do not differentiate among model specifications. I do however only include studies that control for prior student achievement.

The review is broken into three major sections, with several subsections, as follows:

- College Experience
    - Bachelor's Degree & Specific Coursework
    - Master's Degree
    - Selectivity of the University
- Certification
    - Licensure Exam Scores
    - Certification Subject
    - Certification Route
    - National Board Certification
- Experience

College Experience

Colleges play an important role in training teachers. In every state, teachers are required to have a bachelor's degree. To earn a standard teaching license, most states require teachers to earn an education degree before becoming a teacher. With the expansion of alternative routes to the classroom, an increasing number of educators are earning degrees in fields other than teaching, especially secondary teachers of math and science. Not only are alternatively certified teachers entering the classroom with content area degrees, some education schools, like the College of Education and Health Professions at the University of Arkansas, have moved to a Master of Arts in Teaching model, whereby prospective teachers earn a bachelor's degree in a content area and a master's degree in teaching, prior to earning a teaching license.

Graduates of MAT programs enter the profession with an advanced degree under their belt which typically leads to increased pay. Indeed, most traditional public schools pay teachers

based on a step and lane system where the only way to earn a raise is by earning additional college credit toward advanced degrees or with each additional year of experience. Understandably, this has led many teachers to earn a master's degree. In 2011, approximately 43% of teachers nationwide had master's degrees in education; another 12 percent had master's degrees in a field outside of education (Feistritzer, 2011). Whether specific degrees or advanced degrees improve the quality of a teacher is an empirical question and many scholars have sought to answer that question.

Another empirical question in terms of college preparation of teachers is whether the quality of the institution is correlated with the quality of the teachers produced by the college or university. This relationship might occur through one of two routes: higher quality institutions might, through the training provided, somehow increase teachers' abilities more than other schools or higher quality institutions might simply attract higher ability individuals at the onset. After all, teachers earn degrees at a variety of institutions, from the most selective private universities to regional state universities. Below I summarize prior research on the effect of specific degree or coursework, master's degrees, and selectivity of undergraduate institutions on student achievement.

Bachelor's Degree & Subject Specific Coursework

Education degrees

Though most teachers earn degrees from colleges of education, individuals do enter the profession with a myriad of different degrees. Traditionally licensed teachers may earn a degree from a college of education, or sometimes from another academic program as part of an approved teacher training program, the latter being more common for secondary teachers. Alternatively licensed teachers typically earn their bachelor's degree outside of a college of

education. When degrees are earned in conjunction with a college of education, prospective teachers take courses on education theory and practice. Thus, traditionally trained teachers are more likely to have specific education training, as documented by Glazerman, Mayer, and Decker (2006) when examining preparation of Teach for America teachers.

Though traditional education programs require education coursework and similarly many alternative programs require some specific education training, studies on the effectiveness of education-specific training generally reveal findings of no significant effect. In a study of roughly 400 new teachers of grades four through eight in New York City, Rockoff, Jacob, Kane, and Staiger (2008) found no significant differences between teachers with education degrees and those without education degrees in terms of value-added to student achievement. At the high school level, Goldhaber and Brewer (2000) examined the National Educational Longitudinal Study of 1988 (NELS) and found a bachelor's in education had no effect for science teachers and a significant, negative relationship in math. However, it should be noted that Darling-Hammond, Berry, and Thoreson (2001) vigorously criticized Goldhaber and Brewer's methodology and findings as it relates here and to other parts of this review. The criticisms of Darling-Hammond, Berry, and Thoreson are myriad, but they particularly take issue with Goldhaber and Brewer's methodology regarding sample weights and multicolinearity.

Only one study found positive results for education degrees, but even this finding was mixed. Croninger, Rice, Rathbun, & Nishio (2003) found elementary education degrees to have a significant, positive relationship with elementary reading in first grade, but not for math. It seems fitting that the only significant positive result for bachelor's degrees in education happens to be at the earliest level in reading. If anywhere, it seems teachers of reading at the elementary level would benefit from the instruction they receive in a college of education.

Education Coursework

As noted above, many alternative licensure programs require prospective teachers to take some specific education coursework. The Institute for Education Sciences conducted *An Evaluation of Teachers Trained through Different Routes to Certification* (Constantine et al., 2009). As part of the study, the researchers examined the impact of education coursework on student achievement. The study included schools that hired an alternatively certified teacher and had a novice traditionally certified teacher of the same grade and subject, a total of 63 schools in 20 districts. Approximately 2,600 students were randomly assigned to a teacher in these schools. The authors concluded there was not a significant difference in terms of effectiveness between traditionally trained teachers with high levels of coursework and alternatively certified teachers with relatively low levels of education coursework; furthermore, alternatively certified teachers with high levels of teacher training were no more effective than alternatively certified teachers with low levels of teacher training. Similarly, in Florida, Harris and Sass (2011) concluded, "We find almost no evidence that specific undergraduate coursework in education affects an individual's later productivity as a teacher" (p. 811). Their analysis included grades three through ten.

Increasing content knowledge

While specific education training or degrees in education do not have clear benefits for all subject areas, there is some evidence that content knowledge is important for teachers to have. Thus, we would expect teachers with a degree in the subject they are teaching to be more effective educators than those who do not. In fact, the National Math Panel concluded math content knowledge is of the utmost importance for teachers of mathematics (U.S. Department of Education, 2008). Indeed, there is some evidence to support this claim, especially in math. In

several studies, Goldhaber and Brewer (1997a ; 1997b; 2000) conclude secondary teachers with degrees in math are significantly more effective than secondary teachers who did not have a degree in math. Similarly, Harris and Sass (2011) report the number of subject content credits a teacher has in math is significantly positively related to his or her performance teaching high school mathematics.

Though it makes intuitive sense that teachers with higher levels of content knowledge would be more effective teachers, empirical research does not always find significant effects for teachers with specific degrees; Aaronson, Barrow, & Sander (2007) and Rockoff et al. (2008) found no significant differences between teachers with a degree in math or science and teachers who did not have a degree in their field. Yet no study has found a significant negative effect for teachers with a degree in the field they are teaching.

Master's Degree

Using the *High School and Beyond* data, Ehrenberg and Brewer (1994) were the first to control for "sample selection bias and to control for the possible endogeneity of school and teacher characteristics when estimating "gain score" equations" (p. 14). Sample selection bias occurs when the sample in the study is significantly different than the rest of the population. In the case of the *High School and Beyond* data, differences in teacher quality and student characteristics may not be randomly determined and may be a result of sample selection bias. To account for this, Ehrenberg and Brewer used an instrumental variable approach which is a method often used to overcome problems of endogeneity. When looking at the percent of teachers with at least a master's degree, they found mixed results in terms of impact on student achievement; which indicates little or no gain from earning an advanced degree. Since then, the research on master's degrees has been consistent. Indeed, one of the most robust findings in the

teacher quality literature is that advanced degrees typically do not yield increased learning gains for students (Aaronson, Barrow, & Sander 2007; Clotfelter, Ladd, & Vigdor 2007, 2010; Croninger et al. 2003; Hanushek et al. 2005; Krieg 2006; Rockoff et al. 2008; Xu, Hannaway, & Taylor, 2011). In fact, some have found teachers with master's degrees to be significantly worse than those without (Clotfelter, Ladd, & Vigdor 2006).

Even when some positive results for master's degrees are found, they tend to be small or inconsistent between subjects or model specifications (Goldhaber & Anthony, 2007; Cavalluzzo, 2004; Dee 2004; Ferguson & Ladd, 1996; Harris & Sass, 2011; Nye, Konstantopoulos, & Hedges, 2004). For instance, Harris and Sass (2011) report positive findings for teachers of middle school math, but significant negative findings for teachers of middle school reading, high school math and reading. Feng and Sass (2010) report a mix of statistically significant positive findings in math, statistically significant negative findings in reading, and insignificant results in both directions for teachers of special education with advanced degrees. Betts, Zau, & Rice (2003) note master's degrees seemed to have an impact in math but not reading at the elementary level. Ferguson and Ladd (1996) find no evidence teachers with master's degrees were more effective in reading, while they were slightly more effective in math. Among teachers in the Miami-Dade school district, teaching ninth and tenth grade math, graduate degrees were a positive and statistically significant indicator of student achievement; the effect size, however, was very small, ranging from 0.002 to 0.027 in the various model specifications (Cavalluzzo, 2004).

It may be the case that teachers spend a significant amount of time and energy working on their coursework, which results in less energy or effort being spent on preparation for teaching; a hypothesis noted by Constantine et al. (2009). To substantiate the claim, Constantine

et al. noted alternatively certified teachers who were completing education coursework in addition to teaching performed significantly worse than alternatively certified teachers who were not taking additional coursework.

One of the limitations of many of these analyses is the collection of data. Many state departments do not differentiate the varying types of master's degrees and simply report whether teacher's have an advanced degree on a binary scale. Thus, we have less evidence on specific master's degrees than we do on advanced degrees as a whole. It is possible some advanced degrees are beneficial, while others are not. For instance, a math master's degree for math teachers may be beneficial (Goldhaber & Brewer 1997b). Exploring the relationship between specific advanced degrees and student achievement is an area where more research is needed.

Selectivity of the University

There is a strong push to get the best and the brightest into the classroom, presumably because smarter people may make better teachers. It is no secret that some universities are much more selective than others; prestigious universities, like Ivy League schools, are able to recruit individuals with much higher scores on college entrance exams than regional state universities. Perhaps teachers from more prestigious universities are more effective teachers because of their latent ability, which is captured in their prior academic achievement.

Researchers have explored this question using two different measures for the quality of universities; *U.S. News and World Report* rankings of universities and *Barron's* ranking of college competitiveness. The results are not definitive, but suggest teachers from more competitive universities are more effective teachers. For instance, Ehrenberg & Brewer (1994) found student achievement increased more readily when students had a teacher from a high ranking university on the *Barron's* list. Similarly, individuals who attend a very competitive

college, according to *Barron's*, tend to increase student achievement significantly more than individuals from the least competitive universities (Clotfelter, Ladd, and Vigdor 2007, 2010). Rather than use the *Barron's* highly competitive ranking as a binary ranking, Rockoff et al. (2008) include all nine rankings in their analysis and conclude college competitiveness is positively correlated and marginally significant with student achievement for teachers in New York City. Using the *U.S. News* rankings, possibly a less precise measure, Aaronson, Barrow, and Sander (2007) find no difference in terms of teacher effectiveness among Chicago teachers. As a whole, the evidence is very modest, but it suggests that individuals from more competitive universities tend to be more slightly more effective than individuals from less competitive universities.

Certification

Every state requires some form of certification for teachers, with slight variations in the specific qualifications necessary for obtaining a teaching license. In most states, there are a multitude of certifications offered to prospective teachers. Teachers can earn these certifications via a traditional or alternative route, but typically have to pass several exams. In addition to a teaching certificate, veteran teachers can earn the semi-prestigious National Board Certification. This certification is not required to teach, but typically yields a pay increase for qualifying teachers.

The requirements and pathways for certification provide several questions for empirical research. Since teachers can be certified via various routes and in various subjects, researchers can test whether or not teachers are more effective when they are certified in the subject they teach. Researchers can also explore the relationship between licensure exams, which are used to ensure teachers have sufficient content and pedagogical knowledge. Teachers can earn their

certification via a traditional or alternative route. The literature on the relationship between student achievement and licensure exams, certification subject, certification route, and National Board Certification is presented below.

Licensure Exams Scores

Policymakers typically want to ensure teachers have an appropriate level of content knowledge in the subject they wish to teach. Accordingly, tests of basic skills and content knowledge have been required in most states. Additionally, teachers are often required to have a minimum level of pedagogical knowledge. Arkansas is one of many states to use the Praxis test series developed by ETS to assess basic knowledge, content knowledge, and pedagogical knowledge. In fact, the Praxis series is the most widely adopted series of tests used for teacher licensure. Thus, teachers in Arkansas and many other states are required to pass at least three exams: basic skills, pedagogy, and content knowledge.

States use each of these tests on a pass/fail basis. Cut scores, even on the widely used Praxis tests, are determined by policymakers in each state. Consequently, teachers display appropriate levels of knowledge by exceeding the minimum cut score. The logic behind the usage of cut scores implies that an individual who fails the exam by one question is not fit to teach, while the individual who earns a score equal to the cut score is deserving of a teaching certificate. The evidence, however, contradicts this assumption. Using North Carolina data, Goldhaber (2007) examined the impact adjusting cut scores would have on the overall quality of teachers. He concluded North Carolina would eliminate more effective teachers than very ineffective teachers if the state were to adopt Connecticut's higher cut score. He similarly found no improvement in teacher quality when North Carolina increased their cut scores. In Texas,

Hanushek et al. (2005) found no statistical difference between teachers who passed the licensure exam and those who did not.

By using teacher licensure exams, states are able to weed out lower performing individuals, in terms of test performance, but this provides little information to future employers on the ability of prospective teachers. The exam scores may offer more information as a continuous score. There is mounting evidence that licensure exams as a continuous score are somewhat predictive of future performance. In fact, of all the observable teacher characteristics, teacher licensure exams are among the most oft-cited as having a positive relationship with teacher effectiveness. There are two reasons we might think individuals who score higher on licensure exams would make better teachers. First, we expect individuals with a better grasp of their content to be better at instructing students on that topic. At the same time, licensure exams capture not only content knowledge, but intelligence to some extent; and we expect smarter people to be better teachers. There is some evidence to support this claim as both SAT (Boyd et al., 2008) and ACT (Ferguson & Ladd 1996) scores have been found to be positively correlated with teacher effectiveness.

Numerous researchers have found a significant, positive relationship between performance on a teacher licensure exam and a teacher's ability to improve student achievement at the elementary (Clotfelter, Ladd, & Vigdor 2006, 2007; Goldhaber 2007) and high school levels (Clotfelter, Ladd, & Vigdor 2010). Clotfelter, Ladd, and Vigdor (2007) note, "Relative to the estimated effects of class size, the effects of teacher credentials appear to be quite large" (p. 30). The credentials enumerated by Clotfelter, Ladd, and Vigdor include performance on licensure exams.

A common method for estimating the relationship between teacher test scores and student achievement is to standardize all teacher licensure exams. This allows researchers to utilize more data and make comparisons among teachers who have taken a variety of different tests. This provides useful information, but does not provide detailed information on the relationship between specific subject tests and how well teachers perform teaching those subjects. Clotfelter, Ladd, and Vigdor (2010) separate various licensure tests in an effort to explore each subject a bit more closely. They find subject specific test scores in math are significantly, positively correlated with teacher effectiveness. The findings are positive and significant in biology, but are smaller than the findings in math. Interestingly, English teacher licensure exams displayed a significant negative relationship with student achievement in English.

Overall, there is mounting evidence that a teacher's performance on a licensure exam is significantly correlated with their ability to improve student achievement. Yet it is not clear if these findings hold across licensure exams and across teaching subjects. Clearly, more work needs to be done to examine these relationships more thoroughly.

Certification Subject

Common sense would suggest that teachers are better when they are teaching within their certification area. As a matter of fact, there is some evidence to suggest teachers who teach in their field are more effective, especially in math. Clotfelter, Ladd, and Vigdor (2010) note teachers who are certified in their field, both math and English, tend to be more effective teachers. Cavalluzzo (2004) also notes, teachers of mathematics who are certified in their field perform significantly better than teachers who are not certified in their field or are teaching outside of their field.

Certification Route

The standard method for gaining certification is to earn a bachelor's degree in education with a student teaching component. Any other path is considered an alternative route to the classroom. Alternative pathways range in quality and include teachers who enter with emergency certifications or enter through a highly-competitive program, like Teach for America. Alternative certification comes in many varieties, but a typical alternative certification route requires an individual to have a bachelor's degree or relevant experience in the field he or she wishes to teach. These individuals typically must pass the relevant licensure exams and in some cases are required to take specific education coursework prior to or during the first year in the classroom.

The evidence is mixed in terms of traditional and alternative pathways to the classroom. When all non-traditional teachers are lumped together, the results tend to favor traditionally certified teachers. Boyd et al. (2008) note non-certified teachers in New York City were significantly worse than traditionally licensed teachers. Similarly, Clotfelter, Ladd, & Vigdor have found alternatively certified teachers in North Carolina to be significantly worse than traditionally licensed teachers in high school (2010) and in third through fifth grade (2007). Also using North Carolina data, Goldhaber and Anthony (2007) found similar results.

Others have found no significant difference between traditionally and alternatively certified teachers. For instance, Croninger et al. (2003) found no difference in terms of effectiveness for first grade teachers from various certification routes. Similarly, Goldhaber and Brewer (2000) report insignificant differences for 12^{th} grade math and science teachers. The only random assignment study of alternative certification found no significant difference between traditionally and alternatively certified teachers (Constantine et al. 2009).

One problem with the aforementioned studies is how they lump all forms of alternative certification together. As I have noted, alternative licenses can be obtained through a variety of programs with varying levels of difficulty for entrance. When specific programs, which tend to be more selective than colleges of education, are examined the results are a bit more positive for alternative certification. The most highly touted program, Teach for America, was analyzed via a random assignment experiment by Glazerman, Mayer, and Decker (2006). In this analysis, TFA teachers were compared to other teachers within their school regardless of their certification route. Thus, TFA compared to the status quo, rather than TFA compared to only traditionally certified teachers. This is an important distinction, because the schools TFA serves tend to have difficulty hiring only traditionally certified teachers. The results were positive and significant in math, but TFA teachers were no different in terms of increasing student achievement in reading.

TFA has also been analyzed numerous times in non-experimental studies; in these studies, TFA is compared to only traditionally certified teachers. These studies also find a mixture of positive and null results between TFA teachers and traditionally certified teachers. Analyzing data from novice New York City teachers of fourth through eighth grade students, Rockoff et al. (2008) found no statistical difference between TFA teachers and traditionally certified teachers; although, the coefficient for TFA teachers was positive. Similarly, Boyd et al. (2006) and Boyd et al. (2008) do not find significant differences between TFA and traditionally certified teachers in their examination of New York City data for teachers of fourth through eighth grade. Using similar New York City data, Kane, Rockoff, and Staiger (2008) find TFA teachers to be more effective in math instruction, but not so in reading. The most negative findings against TFA come from a 2005 study by Darling-Hammond et al., suggesting uncertified TFA teachers were significantly worse than standard certified teachers. However,

Podgursky (2005) suggests the Darling-Hammond et al. study has severe methodological problems, including the lack of a student fixed effect.

Xu, Hannaway, and Taylor (2011) were the first to estimate the impact of TFA teachers on high school student achievement. Using North Carolina data, they found TFA teachers improved student achievement significantly more than non-TFA teachers in math, science, and English language arts. These findings held even when they compared TFA teachers to more experienced teachers who were teaching within their own field. Taken as a whole, TFA teachers seem to be as effective, if not slightly more effective, in raising student achievement as their traditionally certified counterparts.

The New York City Teaching Fellows (NYCTF) is another selective alternative certification program that has been evaluated numerous times. Like the results for TFA, results for the NYCTF have been mixed, but they tend to be slightly less positive. Rockoff et al. (2008) found NYCTF to be marginally less effective than traditionally certified teachers, but the differences were not significant. Boyd et al (2008) found positive, but insignificant results for NYCTF. Kane et al. (2008) similarly found no significant difference between NYCTF and traditionally certified teachers in math, while the fellows performed slightly worse in reading. As usual, the differences tended to be small.

Another notable alternative certification program is the American Board Certification of Teacher Excellence (ABCTE). Like TFA, ABCTE teachers are certified in a number of different states. Unlike TFA, evaluations of ABCTE have been conducted in only one of the states in which teachers are currently offered initial certification, Florida. As expected, the results are mixed. Using a matching approach with 30 ABCTE teachers of fourth through tenth grade students in Florida, Clark-Tuttle, Anderson, and Glazerman (2009) found ABCTE teachers to be

no different from their matched comparison of traditionally certified teachers in reading, but ABCTE teachers were significantly worse in math. These findings are in contrast to the fixed-effects analyses of Sass (2011), who found ABCTE teachers to be significantly more effective in both math and reading.

Though alternative pathways are not unequivocally better or worse than traditional pathways, there are some clear distinctions between the two routes to the classroom. Alternatively certified teachers are more likely to be from highly competitive universities (Glazerman, Mayer, Decker 2006; Sass 2011; Kane, Rockoff, & Staiger 2008; Boyd et al., 2006). They are also more likely to score higher on the SAT (Sass 2011) or licensure exams (Boyd et al. 2006; Boyd et al. 2008). Indeed, Boyd et al. (2008) note "Only 5% of newly hired Teaching Fellows and TFA teachers in 2003 failed the Liberal Arts and Sciences Test (LAST) exam on their first attempt, while16.2 percent of newly hired traditional teachers failed the LAST exam, and fully 32.5% of uncertified teachers failed the LAST exam" (p. 815). Alternatively certified teachers also are more likely to be male than traditionally certified teachers (Sass 2011).

During the 1999-2000 school year approximately 60% of all newly hired New York City teachers were uncertified (Kane, Rockoff, & Staiger 2008). Alternative routes, like NYCTF and TFA, were created to certify teachers to fulfill the void of certified teachers. The programs have had remarkable success at increasing the number of certified teachers. By 2004-2005 the percentage of uncertified new hires had dropped to only seven percent. At the same time, the number of alternatively certified teachers increased from two percent to 36%. Kane, Rockoff, and Staiger (2008) note the shift from uncertified to alternatively certified teachers was not simply a slight of hand where uncertified teachers were awarded an alternative certificate. The alternative licensure programs attract different individuals, as measured by observable

characteristics. As noted above, teachers in alternative programs tend to be higher performing and have attended more selective universities. This is one of the promises of alternative licensure programs, to attract high performing individuals from other professions. As a result, alternative certification may have the ability to change the composition of the workforce by bringing high-performing individuals and males into the profession. Moreover, alternative certification may also have the added benefit of attracting more minority teachers to the profession (Constantine et al. 2009).

It is important to note that comparisons of traditionally and alternatively certified teachers typically control for other factors, including performance on licensure exams (Boyd et al. 2008, Clotfelter, Ladd, and Vigdor 2007, 2010). Since licensure exams typically have a positive relationship with student achievement gains and alternatively certified teachers typically score significantly better than traditionally certified teachers, the findings for alternative certification may be underestimated.

National Board Certification

In 1987, the National Board for professional Teaching Standards began awarding advanced credentials for teachers. Upon completing the National Board Certification (NBC) procedures, teachers are certified for ten years. Aside from prestige that comes with the certification, some states or school districts offer financial incentives for teachers to complete the certification. Research suggests NBC teachers tend to perform higher in terms of student achievement than those without the certification (Clotfelter, Ladd, & Vigdor 2006, 2007; Golhaber & Anthony 2007; Cavalluzzo, 2004).

The current research, however, cannot causally link NBC certification to increasing student achievement. It is likely the individuals who earn NBC are more productive teachers

prior to attaining the advanced certification. Thus, putting other individuals through the NBC procedures is not likely to improve the overall quality of the teaching workforce. Indeed, the findings of Goldhaber and Anthony (2007) suggest NBC most likely does not add to a teacher's human capital, but actually makes them less effective during the process as they spend extra time completing the required tasks.

Cantrell et al. (2008) conducted an experimental analysis of National Board Certified Teachers. The study consisted of 99 teachers in the Los Angeles area who had applied for National Board Certification and a group of comparison teachers from the same school. Students were randomly assigned to teachers in the study. The authors found no significant difference between NBC teachers and those who had not applied, though the coefficient for NBC teachers tended to be positive. Conversely, teachers who failed to meet the NBC criteria tended to perform worse, in terms of student value-added, although the differences were not significant.

Sanders, Ashton, and Wright (2005) analyzed data from two large North Carolina School districts. Using models that replicate Cavalluzzo (2004) and Goldhaber and Anthony (2007), they found similar effect sizes for NBC teachers. However, when they controlled for the nested structure of students within teacher's classrooms, the results were much less positive. They conclude, "The amount of variability among teachers with the same NBPTS Certification Status is considerably larger than the differences between teachers of different status" (p. 7).

Experience

Teaching is a craft that takes some skill and practice, and it may be expected for teachers to improve with time. The pay system reflects this belief as most teachers earn pay raises with each successive year of experience. This would seem to indicate veteran teachers are more valuable than novice teachers. As a matter of fact, the research is clear that new teachers perform

worse, on average, than more experienced teachers (Boyd et al., 2006; Boyd et al. 2008; Clotfelter, Ladd, & Vigdor 2006, 2007, Croninger et al. 2003; Rockoff 2004; Rivkin, Hanushek, & Kain 2005; Kane, Rockoff, & Staiger 2008; Krieg 2006; Goldhaber & Anthony, 2007; Cavalluzzo, 2004; West & Chingos, 2009; Dee, 2004; Xu, Hannaway, & Taylor, 2011). Hanushek et al. (2005) estimate a first year teacher is, on average, one-half of a standard deviation below the mean.

Teachers tend to perform worse, on average, in their first year and receive large returns in terms of effectiveness from each year of experience. Most of the benefit from additional years of experience are obtained within the first five years of teaching (Boyd et al. 2008; Clotfelter, Ladd, & Vigdor 2007; Rivkin, Hanushek, & Kain 2005). Clotfelter, Ladd, and Vigdor (2007) note more than fifty percent of the gains from experience come within the first two years of experience. After five years, additional years of experience appear to have little benefit in terms of improving teacher effectiveness (Clotfelter, Ladd, & Vigdor 2010, Harris & Sass 2011). For example, Ferguson and Ladd (1996) estimate the benefits of experience in Alabama using a binary variable for teachers with more than five years experience. They report no significant difference between teachers with more or less than five years experience. These findings have been consistent in many settings, including for special education teachers (Feng & Sass 2010).

Rivkin, Hanushek, and Kain (2005) and Boyd et al. (2006) note the large gains from experience in the first few years might be explained by two different phenomena. First, new teachers may actually improve by learning the craft through hands on experience. Another potential explanation is that individuals who feel they are not suited for teaching may leave the teaching workforce. After all, new teachers are much more likely to exit the workforce than are veteran teachers (Hanushek, Kain, & Rivkin 2004). Scholars have reached different conclusions

on these two phenomena. Rivkin, Hanushek, and Kain (2005) conclude "on-the-job learning is the dominant element of the experience effect" (p. 448). This conclusion does not deny there may be some selection effects; indeed Hanushek et al. (2005) note low-performing teachers are more likely to exit in Texas. Boyd et al. (2006) come to the opposite conclusion, "Our teacher fixed effects estimates indicate that the differential effects of experience across pathways is probably due to the differential attrition of weaker teachers across pathways, not to improvements in performance of individual teachers" (p. 212). It seems plausible that both are true, low-performing teachers exit at higher rates and teachers improve with experience in their first few years.

The biggest beneficiaries of experience seem to be alternatively certified teachers. Kane, Rockoff, and Staiger (2008) find alternatively certified teachers improve more rapidly within their first few years of teaching than do traditionally certified teachers. This may be expected, as most alternatively certified teachers have had no prior experience in the classroom, while traditionally certified teachers have had a student teaching experience. Additionally, in many alternative certification programs new teachers are required to complete some coursework within their first two years of teaching. Alternatively certified teachers may improve after these two introductory years as they have fewer distractions in the evenings, like night classes to consume their time. Or it could be the case that the night classes actually improve the quality of the teacher. At the same time, alternatively certified teachers unsuited for teaching may be more likely to exit. Once again, it seems likely the seemingly larger benefits of experience for alternatively certified teachers may come from a little of column A and a little of column B, experience and sorting.

Urban and Rural

Though the research on teacher qualifications is growing, the extant research has been conducted in relatively few places. Much of the research comes from cities, where large databases are easy to build by gaining the cooperation of one school district. These studies, like the one conducted in Chicago by Aaronson, Barrow, and Sander (2007), provide useful information. Few, however, would suggest that Chicago is an accurate reflection of most of the United States. Other studies have been conducted using statewide databases, but few have been conducted in states with a significant rural population.

Take, for example, the studies cited above that examine the impact of certification route on teacher effectiveness. Several of these studies have been conducted in New York City. Just as Chicago is not representative of most of the United States; neither is New York City. Two statewide analyses have been conducted, in North Carolina and Florida. North Carolina has a significant rural population, much more so than Florida. According to the United States Census (2010), 36% of North Carolinians live in rural areas, compared to only 9% in Florida. Though the percent of people living in rural areas is relatively high in North Carolina, it is still significantly lower than in Arkansas; where 45% of the population lives in a rural area.

A number of problems with external validity arise from the small sampling of states and cities. It could be the case that teaching in a city is much different from teaching in a rural school. If so, the characteristics that might make a teacher effective in one location may not be as beneficial in another. It seems unlikely, however, that the characteristics I have discussed here would make teachers more effective in one area than they are in the other. That is to say, scoring high on a licensure exam would seem to be just as beneficial in a city as it would be in a rural

area. Similarly, I can think of no convincing reason to believe experience or pedagogical training would improve performance in one location and not the other.

Table 1: Location of Studies on Teacher Certification Route

Study	Location of Research	Urban/Rural
Boyd et al. (2008)	New York City	Urban
Rockoff et al. (2008)	New York City	Urban
Boyd et al. (2006)	New York City	Urban
Kane et al. (2008)	New York City	Urban
Darling-Hammond et al. (2005)	Houston, Texas	Urban
Clotfelter, Ladd, & Vigdor (2006, 2007, 2010)	North Carolina (Statewide)	Mix
Goldhaber & Anthony (2007)	North Carolina (Statewide)	Mix
Xu, Hannaway, & Taylor (2011)	North Carolina (Statewide)	Mix
Sass (2011)	Florida (Statewide)	Mix
Croninger et al. (2008)	National Sample (Early Childhood Longitudinal Study)	Mix
Goldhaber & Brewer (2000)	National Sample (National Educational Longitudinal Study of 1988)	Mix
Glazerman, Mayer, & Decker (2006)	6 Regions	Mostly urban
Constantine et al. (2009)	7 States (63 schools, 20 districts)	Mostly urban (1 rural district)
Clark-Tuttle, Anderson, & Glazerman (2009)	Florida	NA

The real problem for external validity does not come from the differential impact of skills in various locations; rather, threats to external validity arise from the types of individuals that pool in each of these areas. School districts in New York City, Houston, and other large cities may be more likely to draw higher performing individuals than rural school districts in Arkansas. Moreover, many urban school districts have developed relationships with Teach for America, The New Teacher Project, and other alternative teacher training programs that have a strong track record of attracting smart, hardworking individuals to enter the field of teaching. All

this means that alternative certification programs in cities may have a stronger applicant pool, thus improving the quality of individuals entering from an alternative route.

There is mounting evidence that alternatively certified teachers are indeed more academically capable than their traditionally trained peers. Several studies have documented that alternatively certified teachers are more likely to have attended a highly competitive university than traditionally certified teachers. These findings come from examinations of TFA (Glazerman, Mayer, Decker 2006) teachers in Florida (Sass 2011) and New York City (Kane, Rockoff, & Staiger 2008; Boyd et al. 2008). Similarly, alternatively certified teachers have been found to score higher on standardized tests, including the SAT (Sass 2011) and licensure exams (Boyd et al. 2006; Boyd et al. 2008). Though the existing evidence seems clear that alternatively certified teachers do have higher academic credentials, it is not clear that this finding holds for alternatively certified teachers in Arkansas.

If alternative programs in cities benefit from having a more talented pool of individuals, the findings may be overly generous for alternative certification programs broadly speaking. After all, if the average alternative certification program does not draw significantly higher performing individuals into the profession, then the program is attracting individuals of a similar or lesser academic caliber, who just lack training.

On the other hand, the quality of traditionally trained teachers may vary in urban and rural areas. That is, having a larger pool of applicants may allow urban school districts to hire teachers of higher quality. If that is the case, we would essentially have the opposite problem as before, where the impact of alternatively certified teachers would be underestimated because of the prevalence of higher quality traditionally trained teachers in urban areas.

Of the two scenarios, it seems the former is more likely than the latter. There is ample evidence that alternative programs are attracting high quality individuals to urban areas (Boyd et al. 2008). This is likely the case because individuals who select into teaching tend to have lower levels of prior academic achievement, as Podgursky, Monroe, and Watson (2004) demonstrate in Missouri based on ACT scores of prospective teachers. There is also evidence that urban areas are not flush with traditionally trained teachers. In fact, urban areas often have teacher shortages (Ingersoll 2011). All that is to say, the evidence suggests urban areas are more likely to benefit from alternative programs. The question is whether or not they benefit more than rural areas.

Conclusion

Since the Coleman Report (Coleman et al. 1966) was released, there has been a debate about the impact of money, family life, schools, and teachers on a student's educational achievement. Coleman and his colleagues concluded that family background and socioeconomic status were the most important factors in determining how well a student achieved. The report also noted that schools and teachers have significant effects on student achievement. Since that time, we have come to understand just how important having a quality teacher actually is for students. Kane, Rockoff, and Staiger (2008) note teachers in the top quartile are one-third of a standard deviation better than the bottom quartile. Similarly, Hanushek et al. (2005) conclude student achievement increased by 0.15 standard deviations with a one standard deviation increase in teacher quality. There is no doubt, teachers matter.

Researchers have attempted to identify what, if any, observable teacher characteristics lead to the differential performance of teachers. Studies have examined teacher licensure, years of experience, college degrees, and a host of other observable characteristics. In most areas, the results are mixed. Yet in almost every analysis the differences between two groups tend to be

dwarfed by the differences within each group. That is, there is more variation within traditionally certified teachers as there is between traditionally and alternatively certified teachers. Similarly, there is tremendous variation among teachers with the same level of experience or similar college preparation. Indeed, Aaronson, Barrow, and Sander (2007) note observable teacher characteristics in their analyses "explain at most 10% of the total variation in estimated teacher quality" (p. 97).

The findings are not always significant, yet the results do provide useful information for policymakers. This dissertation will make a unique contribution to the literature by examining the impact of observable teacher characteristics on student achievement in a state that does not closely resemble other markets that have been analyzed.

Chapter III - Methods

In this chapter, I present a description of the data and methods used to answer the following research questions:

(1). Do traditionally certified teachers generate significantly larger increases in student achievement on standardized exams than alternatively certified teachers?

(2). Are teacher licensure exams scores related to gains in student achievement?

(3). Do factors that determine teacher salaries impact student achievement?

(4). Do secondary teachers with a degree in the field they are teaching generate significantly larger learning gains than teachers without a degree in the content field?

The Data

To examine the differences between traditionally and alternatively certified teachers and the relationship of licensure test scores to student achievement, a variety of administrative data on Arkansas public school teachers and students are utilized. These data were provided by the Arkansas Department of Education (ADE).

Teacher Data

The Arkansas Department of Education supplied the data, which were gathered by various agencies within the department, as well as the Department of Higher Education. All of these agencies applied a unique ten digit identification number to each teacher, making it possible to link records from each agency. These data indicate in which school and district a teacher is working in a given year. Additionally, a job code is provided for each teacher. By matching this code to the Statewide Information Systems course code numbering system, I was able to identify the grade and subject of which each individual was the teacher of record.

In allowing me to place a teacher within a subject and grade within a school and district, the data provide useful information regarding each teacher's preparation. When teachers earned their college degree in Arkansas, I am able to identify which college or university they completed their degree. For those individuals who earned their bachelor's outside of the state, the data do not indicate where they attended college. Also included in the data are important descriptive characteristics, including race, gender, whether the teacher has an advanced degree, and whether they are a traditionally or alternatively certified teacher.

The data provide a wealth of information regarding a teacher's performance on state licensure exams. These tests include the Praxis I examinations in mathematics, reading, and writing; as well as Praxis II examinations in content areas and professional knowledge. Many of the more experienced teachers took their exams in the early 1980s as teacher testing was implemented in Arkansas. At the time, Arkansas teachers were required to pass the National Teacher Exam (NTE) provided by Educational Testing Service (ETS). Around 2000, Arkansas switched to the Praxis series of tests also produced by ETS.

To allow us to compare test scores of teachers taking exams in different testing regimes and formats, I standardize all test scores using the cohort of all teachers in our database for whom I have a score for that test code within an appropriate score range. That is, teachers test scores were standardized for the NTE and Praxis exam separately. Bi-modal distributions were detected for some NTE test codes. Since the test codes with bi-modal distributions represent over 12% of all test scores, bi-modal test codes were split into two distributions. The scores were then normalized over the relative subset. Scores were then winsorized to reduce the effects of outliers by replacing outlier variables rather than eliminating observations from the data (Tukey 1962; Winer 1971). The outliers are replaced, rather than eliminated, with the standardized scores at

the 1st percentile and 99th percentile. Since the test scores were normalized over a range of years it implicitly assumes the distribution of scores on a particular test does not change over time, and that the distribution of teacher ability, as measured by test scores, is constant over time.

The aforementioned information on teachers is included for a large portion of the teachers in the data. A much smaller sample of the teachers also have information that indicates their licensure area and their college grade point average. Also appearing in the data to a limited extent are the Classification of Instructional Program (CIP) codes. CIP codes specify the exact degree an individual has received. Other useful, but limited data include teacher scores on their Praxis III evaluation. The Praxis III is an observational evaluation of teaching practices.

Student Data

Like the teacher data, the student data used in this analysis are panel data. Using a unique ten digit student identifier, the data indicate which grade and school a student is enrolled in during a given year. The data also indicate whether the student is enrolled to receive free or reduced price lunches under the National School Lunch Act. Additionally, the data indicate whether the student has an Individualized Education Program (IEP) or is an English language learner. The data also include demographic information on the student's race and gender.

The student data include observations for all students in tested grades between the fiscal years of 2002 and 2010. Test scores appear for students in grades three through eight in math and language arts. These scores come from the state's Benchmark achievement exam. The Benchmark exam is a vertically aligned test; nevertheless test scores are standardized within each grade and year with a mean of zero and a standard deviation of one. For secondary students, the data also include scores on end-of-course exams in algebra and geometry; and the grade 11 literacy examination.

Table 2: Student demographics for last year in sample

Student Demographic	Benchmark	Geometry	Algebra	ELA
Total Students	209,844	29,455	31,585	25,914
Free or reduced price lunch	56.0%	50.8%	54.2%	50.0%
English language learner	5.5%	2.3%	1.6%%	0.0%
Individualized Education Program	10.8%	0.0%	0.7%	0.0%
Female	49.1%	50.4%	49.9%	50.6%
White	67.4%	68.4%	68.0%	69.5%
Black	22.1%	21.4%	20.6%	21.6%
Hispanic	8.2%	7.4%	8.3%	6.0%

Analytic Strategy

Ideally I would be able to match teachers to specific students and know that students are randomly assigned to teachers. However, students are rarely randomly assigned to teachers and with these data, student teacher matching is not possible. This leaves me with two options: either use school level data where the ratio of traditional to alternatively certified teachers and average praxis scores are used as explanatory variables or teacher level data with school value-added scores as the dependent variable. The latter seems more appropriate because that model specification has the variable measured with error as the dependent variable. Under this scenario the coefficients estimates should be unbiased even though the standard errors are higher. To more accurately estimate the standard errors, the error terms are clustered by school and grade.

Value-Added Model – First Stage

Education production functions come in many forms; in this dissertation I used a standard form, as presented by Todd and Wolpin (2003), Sass (2006), and subsequently McGee and Costrell (2010). It is a general cumulative model where student achievement is a function of an individual's cumulative inputs as well as their schools' cumulative contribution to the student's learning. $F_i(t)$ represents the entirety of the individuals personal and family inputs, while $S_i(t)$ represents the input of the school and even the individual teachers.

$$Y_{ijkt} = Y_i[F_i(t), S_i(t), \mu_i, \varepsilon_{it}]. \quad (1)$$

The assumption is made that the cumulative achievement function, Y_{ijkt} , does not vary with time and is additively separable. This allows for the standard production function to be rewritten in an expanded form, where student achievement at time *t* can be expressed as the summation of an individual's input and the summation of a school's input.

$$Y_{ijkt} = \beta \sum_t F_{it} + \gamma \sum_t S_{ijkt} + \varepsilon_{ijkt}. \quad (2)$$

Unfortunately, the data only provide limited information on students and schools; thus an estimation of equation 2 would produce biased results as a result of omitted variables.

To adjust for this, the following empirical model is used to generate unbiased estimates of value-added student achievement.

$$Y_{ijkt} = \alpha \boldsymbol{Y}_{i,t-1} + \beta_1 F + \beta_2 \boldsymbol{S}_{ijkt} + \theta_i + \rho_k + \tau_j + \varepsilon_{ijkt} \quad (3)$$

Following McGee and Costrell (2010) this parsimonious specification is used because it utilizes multiple prior test scores; where. $\boldsymbol{Y}_{i,t-1}$, is a vector of prior student test scores, *F* represents controls for an individual and their family's inputs, school level inputs are controlled for in *S*, the value-added estimate of a grade within a school is captured in τ , while θ, and ρ, represent the unobserved impact of individuals and schools, and ε captures the random error. Ballou, Sanders, and Wright (2004) note that using prior student achievement adequately captures the prior cumulative inputs of an individual and their school. Models of this form have been used by Aaronson (2007) and Hanushek et al. (2005), among others.

The above model includes a student level fixed effect, which I remove in the models estimated for this dissertation. Rather, I use two years of prior test scores in both math and language arts on the right side of the equation to capture student and school time invariant

characteristics. Prior research suggests that these controls adequately account for these inputs (Ballou, Sanders, & Wright 2004). By removing the fixed effects it also increases the statistical power of my estimates. The model used in this estimation is as follows:

$$Y_{Math,t} = \beta_1 Y_{math,t-1} + \beta_2 Y_{math,t-2} + \beta_3 Y_{ELA,t-1} + \beta_4 Y_{ELA,t-2} + \rho + \gamma_{math,j} + \varepsilon \quad (4)$$

$$Y_{ELA,t} = \beta_1 Y_{ELA,t-1} + \beta_2 Y_{ELA,t-2} + \beta_3 Y_{math,t-1} + \beta_4 Y_{math,t-2} + \rho + \gamma_{ELA,j} + \varepsilon \quad (5)$$

Unfortunately, the data provided by the ADE do not allow me to match students directly to teachers. Rather, I am able to match students and teachers to a specific grade in a school. Therefore, value-added estimates are generated at the school-grade level using a random effects estimator. This provides quality estimates at the school-grade level that are normally distributed with a mean of zero. Ideally, I would aggregate by teacher, but the data available do not match students to specific teachers. This, however, does overcome the problem of non-random assignment of students to teachers. The coefficient estimates from these regressions are below.

Table 3: School-by-Grade Value-added Coefficients for the Benchmark Math Exam

	2005-06	2006-07	2007-08
Math Lag 1	0.361***	0.409***	0.415***
	(0.002)	(0.002)	(0.002)
Math Lag 2	0.329***	0.290***	0.301***
	(0.002)	(0.002)	(0.002)
ELA Lag 1	0.138***	0.136***	0.136***
	(0.002)	(0.002)	(0.002)
ELA Lag 2	0.069***	0.071***	0.060***
	(0.002)	(0.002)	(0.002)
Std Dev: Random Effect	0.161	0.154	0.268
	(0.004)	(0.004)	(0.007)
Number of Observations	180,162	299,413	362,031

Table 4 School-by-Grade Value-added Coefficients for the Benchmark ELA Exam

	2006	2007	2008	2009
Math Lag 1	0.072***	0.093***	0.093***	0.094***
	(0.002)	(0.002)	(0.002)	(0.002)
Math Lag 2	0.089***	0.064***	0.046***	0.049***
	(0.002)	(0.002)	(0.002)	(0.002)
ELA Lag 1	0.453***	0.448***	0.455***	0.442***
	(0.002)	(0.002)	(0.002)	(0.002)
ELA Lag 2	0.289***	0.301***	0.304***	0.318***
	(0.002)	(0.002)	(0.002)	(0.002)
Std Dev: Random Effect	0.124	0.111	0.181	0.178
	(0.004)	(0.003)	(0.005)	(0.005)
Number of Observations	180,162	299,413	362,031	368,778

Though end-of-course exams in algebra and geometry and the eleventh grade literacy exam are taken in different years and do not have a similar test taken the year before, the estimation strategy for the high school exams is the same as the strategy outlined above. For each high school exam I use two years of prior math and English language arts scores on the Benchmark exams. These prior tests are typically taken in seventh and eighth grade since geometry is usually taken in 9th or 10th grade and the English language arts exam is taken in eleventh grade. Algebra, on the other hand, is frequently taken in 8th grade. In this case, prior tests may have come from sixth and seventh grade. As the tables below show, the Benchmark exams serve as useful controls for individual and school level time invariant characteristics, providing coefficient estimates similar to those above.

$$Y_{Geo,t} = \beta_1 Y_{math,t-1} + \beta_2 Y_{math,t-2} + \beta_3 Y_{ELA,t-1} + \beta_4 Y_{ELA,t-2} + \rho + \gamma_{Geo,j} + \varepsilon \quad (6)$$

$$Y_{Alg,t} = \beta_1 Y_{math,t-1} + \beta_2 Y_{math,t-2} + \beta_3 Y_{ELA,t-1} + \beta_4 Y_{ELA,t-2} + \rho + \gamma_{Alg,j} + \varepsilon \quad (7)$$

$$Y_{ELA11,t} = \beta_1 Y_{ELA,t-1} + \beta_2 Y_{ELA,t-2} + \beta_3 Y_{math,t-1} + \beta_4 Y_{math,t-2} + \rho + \gamma_{ELA11,j} + \varepsilon \quad (8)$$

Table 5: School-by-Grade Value-added Coefficients for the Geometry EOC

	2008	2009	2010
Math Lag 1	0.403***	0.511***	0.459***
	(0.007)	(0.006)	(0.007)
Math Lag 2	0.408***	0.284***	0.304***
	(0.007)	(0.006)	(0.007)
ELA Lag 1	0.069***	0.086***	0.110***
	(0.008)	(0.007)	(0.007)
ELA Lag 2	0.026***	0.063***	0.002
	(0.007)	(0.007)	(0.007)
Std Dev: Random Effect	0.242	0.203	0.207
	(0.011)	(0.009)	(0.010)
Number of Observations	21,347	28,389	29,455

Table 6: School-by-Grade Value-added Coefficients for the Algebra EOC

	2007	2008	2009	2010
Math Lag 1	0.327***	0.442***	0.454***	0.435***
	(0.007)	(0.006)	(0.006)	(0.007)
Math Lag 2	0.367***	0.211***	0.270***	0.266***
	(0.006)	(0.006)	(0.006)	(0.007)
ELA Lag 1	0.137***	0.163***	0.149***	0.123***
	(0.007)	(0.006)	(0.006)	(0.006)
ELA Lag 2	0.073***	0.084***	0.060***	0.041***
	(0.007)	(0.006)	(0.006)	(0.007)
Std Dev: Random Effect	0.238	0.233	0.204	0.278
	(0.010)	(0.009)	(0.008)	(0.012)
Number of Observations	23,933	30,805	31,322	31,585

Table 7: School-by-Grade Value-added Coefficients for the 11 Grade Literacy Exam

	2009	2010
Math Lag 1	0.097***	0.175***
	(0.007)	(0.006)
Math Lag 2	0.164***	0.059***
	(0.006)	(0.006)
ELA Lag 1	0.362***	0.394***
	(0.007)	(0.007)
ELA Lag 2	0.355***	0.341***
	(0.006)	(0.007)
Std Dev: Random Effect	0.131	0.163
	(0.008)	(0.009)
Number of Observations	25,560	25,914

Regression on Teacher Characteristics - Second Stage

Since I am unable to match students directly to teachers, I am left with two options. I could generate the value-added for a school-grade level using methods similar to above, while including school-grade level averaged information on teachers. That is, I could calculate the ratio of traditional to alternatively certified teachers and average praxis scores and generate estimates of their impact in the value-added estimate. The other option, which I have elected to do, is use teacher level data with school value-added scores as the dependent variable. I chose the latter because under this scenario the coefficient estimates should be unbiased even though the standard errors are higher. I cluster the error terms by school and grade to more accurately estimate the standard errors.

Upon generating the value-added estimates for the school-grade level, I then estimate the following equation:

$$u_{j,k,t} = \alpha + \psi_1 \mathbf{TQ} + \psi_2 \mathbf{TC} + \psi_3 \text{Year} + e_{jk} \quad (3)$$

The dependent variable, $u_{j,k,t}$, is the residual captured as value-added in equations (4) through (8). **TQ** is a vector of observable teacher characteristics typically used to measure teacher quality. This vector includes: experience, experience squared, certification route, Praxis I test scores, Praxis II test scores, and graduate degrees. Teacher characteristics, **TC**, are controlled for in a vector of teacher demographics including race and gender. I use binary indicators for each year. Note that since the dependent variable in this stage is for the entire grade at the school, I see the same dependent variable measure for each teacher within a school grade, and the number of times I see it varies with the number of teachers per grade at the school. For this reason I cluster error terms by school and grade to estimate standard errors more accurately.

Benchmark and High School Analysis

Since I am unable to match teachers to students, I analyze teacher impacts on student achievement in a number of ways to provide a series of robustness checks regarding the research questions. I start with the largest possible sample of teachers and then narrow the sample in subsequent analyses. Using job codes provided by the state, I am able to identify what subject a teacher teaches. The first set of analyses includes all teachers that might reasonably contribute to a student's learning in math or language arts, which I believe includes any core subject teacher. Thus, I include all teachers with a job code for math and language arts; as well as, teachers who taught science and social studies. These can be thought of as all of the core teachers at a given grade in a school (School-Grade). In the second set of analyses I remove teachers who do not teach the specific subject being tested, leaving only teachers with a math job code in the math analyses and language arts in the language arts analyses (School-Grade-Subject). Next, I examine the data while limiting the sample to only new teachers. In this set of analyses and each subsequent set, I conduct the analyses using the School-Grade sample which includes all core

subject teachers. I conduct additional analyses to estimate whether teachers further from the mean in terms of their licensure exam score disproportionately impact student achievement.

For the high school tests, I follow a similar strategy as I have in the Benchmark analyses. First, I estimate the observable teacher characteristics on the entire sample of teachers in each subject. Next, I conduct the analysis while limiting the experience of teachers. At the high school level, there are a number of schools that only have one teacher of algebra, geometry, or eleventh grade English. Therefore, I present the analyses while limiting the data to only these teachers as well.

Hypothesis Tests

Using the methods described above, I will test the following four hypotheses.

- Research Question (1): Do traditionally certified teachers generate significantly larger increases in student achievement on standardized exams than alternatively certified teachers?

  H_0: Alternatively certified teachers increase student achievement at similar rates as traditionally certified teachers.

  H_1: Alternatively certified teachers increase student achievement significantly more or significantly less than traditionally certified teachers.

- Research Question (2). Are teacher licensure exams scores related to gains in student achievement?

  H_0: There is no relationship between a teacher's score on a licensure exam and their effectiveness at improving student achievement.

  H_1: There is a significant relationship between a teacher's score on a licensure exam and their effectiveness at improving student achievement.

- Research Question (3). Do factors that determine teacher salaries impact student achievement?

  H_0: There is no relationship between a teaching experience or graduate degrees and effectiveness at improving student achievement.

  H_1: There is a significant relationship between teaching experience or having a graduate degree on student achievement.

- Research Question (4). Do secondary teachers with a degree in the field they are teaching generate significantly larger learning gains than teachers without a degree in the content field?

  H_0: There is no relationship between having a degree in the content area you are teaching and effectiveness at improving student achievement.

  H_1: There is a significant relationship between having a content degree in the field you are teaching and effectiveness at improving student achievement.

Classical hypothesis tests will be conducted on each of the research questions. For these tests, the coefficients estimated in my regression analyses will be used to determine whether the null hypotheses should be rejected. For each coefficient, *p*-values of less than .05 will be determined as statistically significant. I will note coefficients that have a *p*-value of less than .10, but these marginally significant results will not be used in determining whether the null hypotheses should be rejected.

Summary

The analytic strategy used to answer the research questions in this dissertation requires two steps. In the first, I generate value-added scores for the Benchmark exams, the algebra and geometry end-of-course exams, and the eleventh grade English language arts exam at the school-

grade level. In the second step I make the value-added the dependent variable and I regress it on observable teacher characteristics. These characteristics include the teacher's certification route, scores on licensure exams, and other important information. I estimate a number of different model specifications as a series of robustness checks.

Chapter IV - Results

My analysis of the effect of various teacher characteristics on student value-added achievement is conducted in two ways. Accordingly, this section will be separated into two major sections, benchmark and secondary. The benchmark analyses include student test scores on the benchmark literacy and math exams in grades three through eight. Since I use two years of prior test scores to estimate the value-added score, teacher characteristics for grades three and four are not analyzed in the models.

Using job codes provided by the state, the data allow me to identify what grade and subject a teacher teaches. This allows me to analyze the data in two ways. First, I start with the largest possible sample of teachers, the entire sample of teachers in a given grade in a school, and then I narrow the sample in subsequent analyses. The first set of analyses includes all teachers that might reasonably contribute to a student's learning in math or language arts, which I believe includes any core subject teacher. Thus, I include all teachers with a job code for math and language arts, as well as teachers who taught science and social studies. This level of analyses will be labeled *School-Grade*, since it includes all teachers in a school teaching a specific grade. In the second set of analyses I remove teachers who do not teach the specific subject being tested, leaving only teachers with a math job code in the math analyses and language arts in the language arts analyses. This set of analyses is identified as *School-Grade-Subject*.

Alternatively certified teachers in the sample tend to have less experience than traditionally certified teachers. Therefore, I conduct additional analyses while limiting the sample to all teachers of core subjects in their first five years (*School-Grade-Novice5*) of teaching and in their first ten years of teaching (*School-Grade-Novice10*). In this set of analyses

and each subsequent set, I conduct the analyses using the School-Grade sample which includes all core subject teachers. I conduct a series of additional analyses to estimate whether the teachers further from the mean in terms of their licensure exam score disproportionately impact student achievement.

Benchmark Analyses

Although the data include information on three sections of the Praxis I test: reading, writing, and math, in addition to scores on the Praxis II exam, I find few variables to be significant when I utilize all of the variables in the models because of colinearity. In fact, many of the explanatory variables are highly collinear. Table 8 displays the correlation coefficients between the teacher test data, race, and gender with the probability of independent distributions in parenthesis below the correlation coefficients. I also include a Praxis I variable; which is a composite score created by taking an average of the three Praxis I subtests. The Praxis II test is highly correlated with each section of the Praxis I and even more so with the Praxis I composite score. This may be expected as people who are knowledgeable and/or test well would be expected to perform well on all Praxis I and II exams. Notably, race is also significantly correlated with each test. As we know from national and Arkansas education data, there is an achievement gap between white and minority students and this is displayed by the disparate average performance of white and minority teachers in Arkansas. Also, close to 90% of teachers in the sample are white and when I place a race indicator in the model it is significantly related to performance in the classroom due to correlation with our other variables. Thus, I remove the race indicator from the models.

Table 8: Correlation matrix using least restrictive school-grade math sample

	Praxis I	Praxis II	Praxis I-Math	Praxis I-Reading	Praxis I-Writing	White
Praxis I	1.0000					
Praxis II	0.6083 (0.0000)	1.0000				
Praxis I-Math	0.8095 (0.0000)	0.4269 (0.0000)	1.000			
Praxis I-Reading	0.8358 (0.0000)	0.5825 (0.0000)	0.5238 (0.0000)	1.0000		
Praxis I-Writing	0.7847 (0.0000)	0.4732 (0.0000)	0.4151 (0.0000)	0.5147 (0.0000)	1.0000	
White	0.2201 (0.0000)	0.2340 (0.0000)	0.1891 (0.0000)	0.1907 (0.0000)	0.1537 (0.0000)	1.0000

School-Grade Analyses

Table 9 provides descriptive information about the teachers in the sample. For the information in the tables I use all teachers in the least restrictive math model, which includes teacher's years of experience and performance on the Praxis II. The standardized z-scores for the tests do not equal zero because this is a subsample of our entire population of test takers. The table displays the differences between traditionally and alternatively certified teachers in the data. I test the significance of these differences in every domain, except grade level. I note alternatively certified teachers tend to score higher on all licensure exams, are more likely to be male, less likely to have a graduate degree, and are more novice than traditionally certified teachers in our sample.

As mentioned above, the first set of analyses attribute the value-added to all teachers of core subjects at a specific grade within a school. I display the regression results for the school-grade level math analyses in Table 10 and English language arts (ELA) in Table 11. In each

analysis, I cluster error terms at the school-grade level. The sample sizes range from 17,196 teachers in the least restrictive sample to 5,567 in both math and ELA.

Table 9. Teacher demographics by licensure route from least restrictive math model

Teacher Characteristic	Non-Traditional	Traditional
Number of Grade 5 Teachers	94	3,789
Number of Grade 6 Teachers	136	4,011
Number of Grade 7 Teachers	319	4,132
Number of Grade 8 Teachers	349	4,366
Teachers with a Graduate Degree	26.3%***	39.0%
Avg. Praxis I-Math z-score	.4778 (.8442)***	-.0621 (9667)
Avg. Praxis I- Writing z-score	.2630 (.8081)***	-.0493 (.8406)
Avg. Praxis I- Reading z-score	.3765 (.6511)***	-.0097 (.8738)
Avg. Praxis II- Prof. Knowledge z-score	.1587 (.7673)***	-.0327 (.8117)
White	88.1%	88.2%
Black	9.7%	11.0%
Female	66.0%***	81.5%
Years of Experience	1.8 (2.1)***	12.9 (9.7)

Table 10: Math school-grade level results with clustered error terms

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Alt. Certification	-0.0015	-0.0109	-0.0177*	-0.0206**	-0.0268**	-0.0260**
	[0.0066]	[0.0088]	[0.0105]	[0.0102]	[0.0117]	[0.0128]
Alt. Cert/Exp. Interaction Term		0.0056**	0.0053*	0.0048*	0.0058*	0.0058*
		[0.0026]	[0.0029]	[0.0029]	[0.0031]	[0.0031]
Alt. Cert/Grad. Degree Interaction						-0.0037
						[0.0199]
Praxis II-Prof.				0.0126***	0.0154***	0.0154***
				[0.0025]	[0.0048]	[0.0048]
Praxis I- Avg.			0.0104***		0.0002	0.0002
			[0.0039]		[0.0049]	[0.0049]
BA Not earned in Arkansas		-0.0029	0.0004	-0.0086	-0.0069	-0.0070
		[0.0052]	[0.0139]	[0.0060]	[0.0143]	[0.0143]
Graduate Degree		0.0005	-0.0069	-0.0028	-0.0088	-0.0085
		[0.0039]	[0.0059]	[0.0044]	[0.0060]	[0.0063]
Experience		0.0011	0.0034*	0.0010	0.0025	0.0025
		[0.0007]	[0.0020]	[0.0007]	[0.0021]	[0.0021]
Experience Squared (units of 100)		-0.0050**	-0.0214	-0.0044*	-0.0161	-0.0161
		[0.0021]	[0.0148]	[0.0025]	[0.0153]	[0.0153]
2007		0.0141**	0.0121*	0.0131**	0.0117*	0.0117*
		[0.0056]	[0.0065]	[0.0055]	[0.0064]	[0.0064]
2008		0.0167**	0.0124*	0.0161**	0.0124*	0.0124*
		[0.0066]	[0.0075]	[0.0063]	[0.0073]	[0.0073]
Female		0.0062	0.0151**	0.0077	0.0107*	0.0106*
		[0.0044]	[0.0061]	[0.0049]	[0.0063]	[0.0063]
Constant	-0.0003	-0.0151*	-0.0209**	-0.0136*	-0.0142	-0.0142
	[0.0043]	[0.0079]	[0.0101]	[0.0078]	[0.0104]	[0.0104]
Observations	20,009	17,196	5,916	13,407	5,567	5,567
R-squared	0.0000	0.005	0.008	0.009	0.012	0.012

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 11: Language arts school-grade level results with clustered error terms

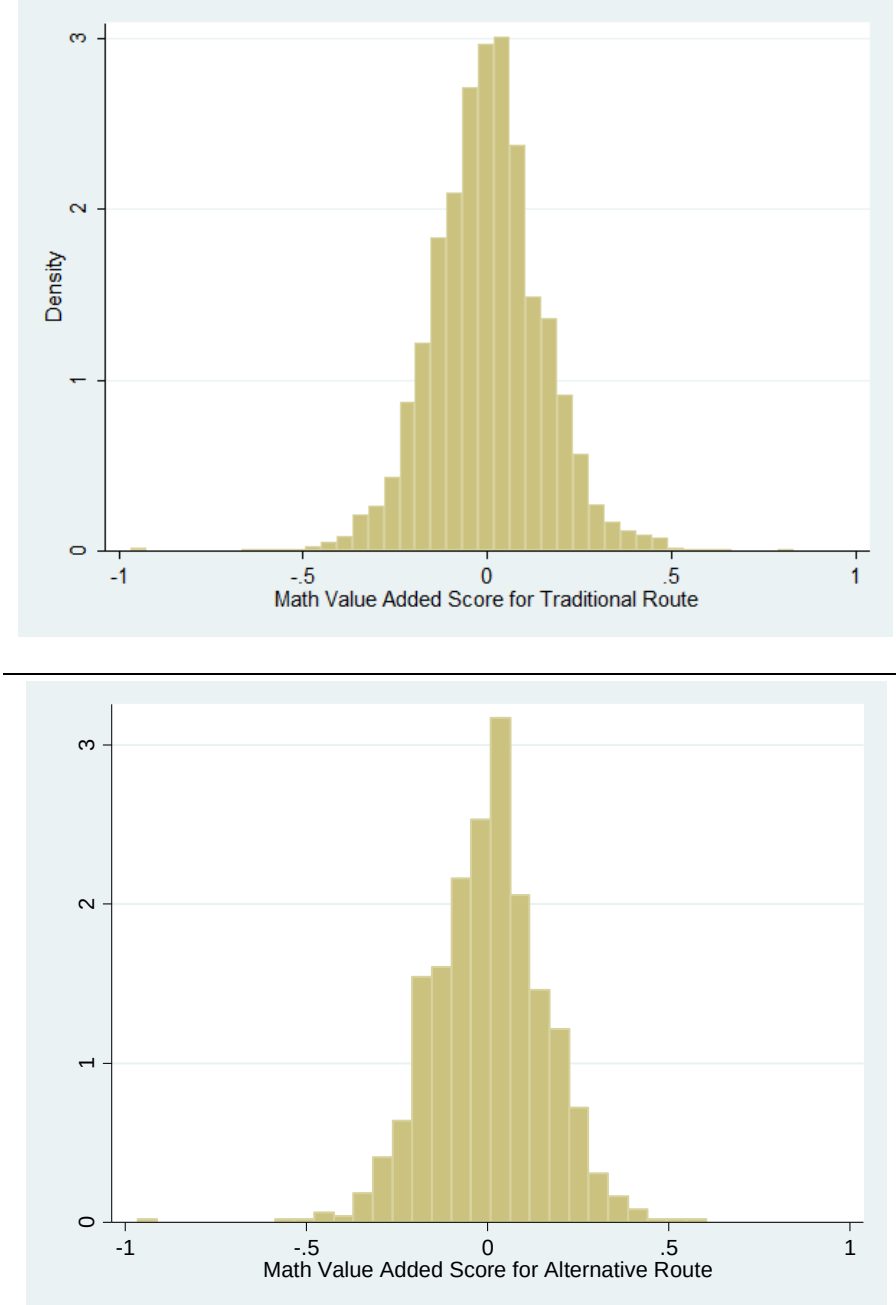
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Alternative Certification	-0.0079	-0.0088	-0.0174**	-0.0138*	-0.0187**	-0.0126
	[0.0051]	[0.0067]	[0.0076]	[0.0076]	[0.0085]	[0.0091]
Alt. Cert/Exp. Interaction Term		0.0041**	0.0053**	0.0044**	0.0049**	0.0053**
		[0.0021]	[0.0022]	[0.0022]	[0.0023]	[0.0024]
Alt. Cert/Grad. Degree Interaction						-0.0297*
						[0.0161]
Praxis II				0.0072***	0.0064*	0.0068*
				[0.0019]	[0.0037]	[0.0037]
Praxis I			0.0076***		0.0039	0.0036
			[0.0028]		[0.0036]	[0.0037]
BA Not earned in Arkansas		-0.0003	0.0077	-0.0024	0.0040	0.0030
		[0.0038]	[0.0099]	[0.0045]	[0.0109]	[0.0109]
Graduate Degree		0.0013	-0.0042	3.07e-05	-0.0057	-0.0032
		[0.0027]	[0.0042]	[0.0031]	[0.0043]	[0.0046]
Experience		0.0009**	0.0008	0.0006	0.0003	0.0003
		[0.0005]	[0.0012]	[0.0005]	[0.0013]	[0.0013]
Experience Squared (units of 100)		-0.0027*	-0.0011	-0.0008	0.0021	0.0021
		[0.0014]	[0.0085]	[0.0018]	[0.0089]	[0.0090]
2007		0.0036	0.0012	0.0020	0.0005	0.0003
		[0.0037]	[0.0043]	[0.0038]	[0.0043]	[0.0043]
2008		-0.0005	-0.0008	-0.0021	-0.0007	-0.0008
		[0.0045]	[0.0053]	[0.0045]	[0.0053]	[0.0053]
Female		0.0077**	0.0092**	0.0036	0.0059	0.0053
		[0.0034]	[0.0045]	[0.0036]	[0.0047]	[0.0047]
Constant	-0.0000	-0.0106*	-0.0094	-0.0048	-0.0051	-0.0055
	[0.0030]	[0.0056]	[0.0070]	[0.0057]	[0.0072]	[0.0073]
Observations	20,009	17,196	5,916	13,407	5,567	5,567
R-squared	0.0003	0.002	0.007	0.005	0.008	0.009

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

At first glance alternatively certified teachers seem to be significantly lower performing than traditionally certified teachers. In each model in both math and ELA, the coefficient on alternative certification is negative; however, the coefficient is only significant when the teacher's performance on licensure exams is controlled for in the model. It is clear from these analyses that teachers who score higher on licensure exams tend to perform better in terms of raising student achievement. Thus, controlling for licensure scores is inappropriate since licensure exam scores are significantly correlated with licensure route and alternatively certified teachers in this sample score significantly higher on licensure exams. Indeed, one of the promises of alternative certification is the potential of drawing high performing individuals into the classroom.

On average, teachers from both routes perform similarly. However, there is significant variation in the quality of teachers within both the traditional and alternative paths. As Figure 1 displays, both groups of teachers are fairly normally distributed, with the mean around zero. In both cases, there are a number of teachers who are significantly above or below the mean in terms of value added scores.

Figure 1: Distribution of VAM Scores for Traditionally/Alternatively Certified Teachers



The spread for the two groups appears to be slightly wider for alternatively certified teachers. There is primarily one reason for this occurrence. There are fewer observations of alternatively certified teachers, 898 compared to 16,298 for traditionally certified teachers (see

Table 12). For both groups, the mean is 0.002 and the standard deviations are similar, 0.151 for traditionally certified teachers and 0.157 for alternatively certified teachers.

Table 12: Summary Statistics of Traditional/Alternative Math VAM Scores

	Observations	Mean	Standard Deviation	Minimum VAM Score	Maximum VAM Score
Traditional	16,298	0.002	0.151	-0.966	0.834
Alternative	898	0.002	0.157	-0.966	0.604

Though I find little evidence of increased effectiveness from additional experience as a whole, it is marginally significant in one model in math and significant in one ELA model, I do find alternatively certified teachers seem to improve more rapidly. I include an interaction term for alternative certification and years of experience. The coefficient on the interaction term is positive and statistically significant or marginally significant in each model in both subject areas. Most alternatively certified teachers enter the classroom with little to no experience actually teaching, while traditionally certified teachers have completed a student teaching experience prior to entering the classroom as a teacher of record. Thus, it may be expected that alternatively certified teachers would benefit more from an additional year of experience.

School-Grade-Subject Analyses

In order to get a more direct match of teachers to students in the tested areas, I limit the teachers in the sample to only those listed as a math or language arts teacher. This considerably lowers the number of teachers in the analysis, but potentially provides a more accurate assessment of the teacher's impact on student achievement (See Tables 13 and 14) As before, I find a significant positive relationship between a teacher's performance on the Praxis II professional knowledge exam and student achievement in math ($p. <0.01$); however, the exam score does not have a statistically significant effect on ELA scores. When I restrict the sample to

only teachers of the tested subjects, alternatively certified teachers are not statistically different from traditionally certified teachers in their effects on student value-added in any of the models. Interestingly, the interaction term between alternatively certified teachers and graduate degrees is negative and marginally significant in English Language Arts.

Table 13: Math school-grade-subject level results with clustered error terms

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Alt. Certification	-0.0054	-0.0106	0.0061	-0.0197	-0.0019	-0.0070
	[0.114]	[0.0161]	[0.0182]	[0.0183]	[0.0199]	[0.0205]
Alt. Cert/Exp. Interaction		0.0036	0.0005	0.0065	0.0037	0.0029
		[0.0042]	[0.0048]	[0.00430]	[0.0048]	[0.0046]
Alt. Cert/Grad. Degree Interaction						0.0305
						[0.0394]
Praxis II				0.0172***	0.0182**	0.0177**
				[0.0045]	[0.0089]	[0.0090]
Praxis I			0.0134**		0.0017	0.0023
			[0.0068]		[0.0087]	[0.0088]
BA Not earned in AR		-0.0153*	-0.0403*	-0.0203**	-0.0393	-0.0390
		[0.0081]	[0.0226]	[0.0104]	[0.0243]	[0.0241]
Graduate Degree		0.0114	0.0029	0.0089	0.0021	-0.0005
		[0.0074]	[0.0107]	[0.0083]	[0.0109]	[0.0112]
Experience		-0.0008	0.0073**	-0.0005	0.0068*	0.0068*
		[0.0011]	[0.0037]	[0.0013]	[0.0037]	[0.0037]
Experience Squared (units of 100)		-0.0009	-0.0370	-0.0022	-0.0353	-0.0351
		[0.0034]	[0.0277]	[0.0043]	[0.0280]	[0.0280]
2007		0.0202***	0.0181*	0.0229***	0.0179*	0.0179*
		[0.0076]	[0.0100]	[0.0082]	[0.0100]	[0.0100]
2008		0.0238***	0.0224**	0.0252***	0.0197*	0.0197*
		[0.0081]	[0.0106]	[0.0086]	[0.0108]	[0.0108]
Female		0.0413***	0.0424***	0.0452***	0.0355***	0.0364***
		[0.0075]	[0.0113]	[0.0089]	[0.0120]	[0.0121]
Constant	0.0064	-0.0304***	-0.0531***	-0.0347***	-0.0447**	-0.0445**
	[0.0050]	[0.0109]	[0.0170]	[0.0124]	[0.0178]	[0.0177]
Observations	5,512	4,620	1,669	3,704	1,591	1,591
R-squared	0.000	0.019	0.029	0.029	0.033	0.034

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 14: Language arts school-grade-subject level results with clustered error terms

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Alternative Certification	-0.0122	-0.0206	-0.0252	-0.0258	-0.0285	-0.0145
	[0.0097]	[0.0145]	[0.0159]	[0.0168]	[0.0191]	[0.0205]
Alt. Cert/Exp. Interaction Term		0.0040	0.0057	0.0006	0.0017	0.0015
		[0.0058]	[0.0065]	[0.0069]	[0.0072]	[0.0072]
Alt. Cert/Grad. Degree Interaction Term						-0.0661*
						[0.0361]
Praxis II				0.0048	-0.0034	-0.0034
				[0.0029]	[0.0067]	[0.0066]
Praxis I			0.0074		0.0091	0.0088
			[0.0051]		[0.0065]	[0.0065]
BA Not earned in Arkansas		0.0050	0.0060	0.0050	-0.0184	-0.0194
		[0.0053]	[0.0175]	[0.0063]	[0.0181]	[0.0187]
Graduate Degree		0.0008	0.0002	0.0011	-0.0037	-0.0003
		[0.0045]	[0.0074]	[0.0053]	[0.0077]	[0.0078]
Experience		7.23e-05	0.0013	-0.0002	0.0003	0.0004
		[0.0007]	[0.0023]	[0.0009]	[0.0022]	[0.0022]
Experience Squared (units of 100)		-0.0006	-0.0022	0.0018	0.0032	0.0029
		[0.0022]	[0.0161]	[0.0031]	[0.0157]	[0.0157]
2007		0.0067	0.0065	0.0067	0.0064	0.0063
		[0.0055]	[0.0073]	[0.0058]	[0.0075]	[0.0075]
2008		0.0075	0.0097	0.0072	0.0119	0.0120
		[0.0060]	[0.0083]	[0.0063]	[0.0083]	[0.0082]
Female		0.0141*	0.0192	0.0152	0.0264*	0.0260*
		[0.0080]	[0.0130]	[0.0093]	[0.0139]	[0.0140]
Constant	0.0028	-0.0140	-0.0273*	-0.0157	-0.0293*	-0.0304**
	[0.0034]	[0.0102]	[0.0145]	[0.0111]	[0.0151]	[0.0151]
Observations	6,707	5,789	1,874	4,523	1,737	1,737
R-squared	0.001	0.003	0.010	0.005	0.014	0.017

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

School-Grade-Novice Analyses

As I noted in Table 9, traditionally certified teachers in our sample have much more experience than non-traditionally licensed teachers; an average of 12.9 years compared to 1.8 years. I conduct additional analyses to take into account the varying levels of experience. Essentially, I want to compare new non-traditional teachers to new traditional teachers. I do this by limiting the sample to teachers within their first five years of experience. Once again, the results for math and language arts are for the most part consistent with the above analyses (See Tables 15 and 16). Praxis II scores remain the strongest predictor of performance in both subjects. The coefficients on alternative certification remain negative in each model, but are only marginally significant in one math model and significant ($p. <0.05$) in two ELA models. Again, the difference between traditionally and alternatively certified teachers is only significant when licensure exam scores are included in the model.

Table 15: Math school-grade level results for novice teachers with clustered error terms

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Alternative Certification	-0.0037	-0.0011	-0.0144	-0.0128	-0.0247*	-0.0227
	[0.0074]	[0.0108]	[0.0122]	[0.0117]	[0.0129]	[0.0139]
Alt. Cert/Exp. Interaction Term		-0.0009	0.0007	-0.0022	0.0022	0.0024
		[0.0051]	[0.0055]	[0.0053]	[0.0058]	[0.0058]
Alt. Cert/Grad. Degree Interaction Term						-0.0092
						[0.0213]
Praxis II				0.0141***	0.0197***	0.0199***
				[0.0042]	[0.0062]	[0.0062]
Praxis I			0.0178***		0.0067	0.0066
			[0.0055]		[0.0065]	[0.0065]
BA Not earned in Arkansas		0.0062	0.0033	0.0095	-0.0033	-0.0037
		[0.0078]	[0.0169]	[0.0099]	[0.0180]	[0.0180]
Graduate Degree		0.0007	-0.0040	-0.0028	-0.0083	-0.0068
		[0.0063]	[0.0078]	[0.0066]	[0.0079]	[0.0087]
Experience		0.0077	0.0045	0.0038	-0.0019	-0.0019
		[0.0053]	[0.0062]	[0.0055]	[0.0065]	[0.0065]
Experience Squared (units of 100)		-0.0011	-0.0007	-0.0006	0.0002	0.0002
		[0.0010]	[0.0013]	[0.0010]	[0.0013]	[0.0013]
2007		0.0173**	0.0182**	0.0155**	0.0172*	0.0172*
		[0.0075]	[0.0089]	[0.0075]	[0.0088]	[0.0088]
2008		0.0208**	0.0195**	0.0208**	0.0195**	0.0195**
		[0.0086]	[0.0099]	[0.0085]	[0.0097]	[0.0097]
Female		0.0053	0.0094	0.0039	0.0051	0.0048
		[0.0065]	[0.0081]	[0.0075]	[0.0087]	[0.0088]
Constant	0.0034	-0.0253**	-0.0232*	-0.0176	-0.0111	-0.0113
	[0.0050]	[0.0109]	[0.0125]	[0.0116]	[0.0132]	[0.0132]
Observations	6,012	5,819	3,143	4,791	2,825	2,825
R-squared	0.000	0.005	0.010	0.010	0.019	0.019

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 16: Lit school-grade level results for novice teachers with clustered error terms

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Alternative Certification	-0.0058	-0.0085	-0.0216**	-0.0169*	-0.0253**	-0.0179*
	[0.0055]	[0.0081]	[0.0093]	[0.0097]	[0.0105]	[0.0107]
Alt. Cert/Exp. Interaction Term		0.0033	0.0071	0.0055	0.0078	0.0082
		[0.0040]	[0.0047]	[0.0046]	[0.0051]	[0.0051]
Alt. Cert/Grad. Degree Interaction Term						-0.0339*
						[0.0175]
Praxis II				0.0038	0.0036	0.0043
				[0.0029]	[0.0047]	[0.0047]
Praxis I			0.0095**		0.0083*	0.0077
			[0.0038]		[0.0047]	[0.0047]
BA Not earned in Arkansas		-0.0031	0.0077	0.0048	0.0031	0.0016
		[0.0056]	[0.0123]	[0.0076]	[0.0141]	[0.0140]
Graduate Degree		0.0031	-0.0036	0.0010	-0.0064	-0.0011
		[0.0043]	[0.0057]	[0.0048]	[0.0061]	[0.0066]
Experience		0.0043	0.0037	0.0022	-0.0008	-0.0009
		[0.0036]	[0.0042]	[0.0039]	[0.0045]	[0.0045]
Experience Squared (units of 100)		-0.0006	-0.0007	-0.0003	0.0000	0.0001
		[0.0007]	[0.0008]	[0.0007]	[0.0009]	[0.0009]
2007		0.0037	0.0042	0.0022	0.0031	0.0028
		[0.0046]	[0.0056]	[0.0048]	[0.0058]	[0.0058]
2008		-0.0032	0.0018	-0.0048	0.0014	0.0013
		[0.0056]	[0.0065]	[0.0057]	[0.0067]	[0.0067]
Female		0.0044	0.0044	0.0002	0.0014	0.0005
		[0.0049]	[0.0063]	[0.0056]	[0.0067]	[0.0067]
Constant	0.0000	-0.0101	-0.0088	-0.0018	-0.0007	-0.0016
	[0.0034]	[0.0078]	[0.0091]	[0.0080]	[0.0098]	[0.0099]
Observations	6,012	5,819	3,143	4,791	2,825	2,825
R-squared	0.000	0.003	0.008	0.004	0.009	0.011

Standard errors in brackets

*** p<0.01, ** p<0.05, * p<0.1

Teachers in the Tails

The above analyses suggest teacher performance on licensure exams, particularly the Praxis II content knowledge exam, is significantly correlated with student achievement. It could be the case that teachers in the tails of the Praxis distributions have a disproportionate impact on student achievement. To test this, I create indicator variables for teachers in the top and bottom ten percent and top and bottom 25 percent in terms of performance on the Praxis I and II. I then replicate the previously-presented analyses with indicators for teachers with Praxis I or II score in top or bottom of the distribution. As expected, higher scoring teachers tend to be significantly more effective at increasing student achievement in both subjects; although statistical significance is detected more readily in math.

Table 17: Math testing tails, school-grade level results with clustered error terms

Variables	Praxis I-10%	Praxis I-25%	Praxis II-10%	Praxis II-25%
Alternative Certification	-0.0172	-0.0169	-0.0201**	-0.0201**
	[0.0105]	[0.0105]	[0.0102]	[0.0102]
Alt. Cert/Exp. Interaction Term	0.0054*	0.0054*	0.0045	0.0050*
	[0.0029]	[0.0029]	[0.0029]	[0.0029]
Licensure Test-Top	0.0183*	0.0132**	0.0169***	0.0080*
	[0.0099]	[0.0066]	[0.0063]	[0.0043]
Licensure Test-Bottom	-0.0176**	-0.0013	-0.0274***	-0.0157***
	[0.0079]	[0.0066]	[0.0060]	[0.0045]
BA Not earned in Arkansas	-0.0003	0.0008	-0.0078	-0.0077
	[0.0137]	[0.0139]	[0.0060]	[0.0060]
Graduate Degree	-0.0066	-0.0066	-0.0023	-0.0025
	[0.0059]	[0.0059]	[0.0043]	[0.0043]
Experience	0.0033*	0.0032	0.0010	0.0010
	[0.0020]	[0.0020]	[0.0007]	[0.0007]
Experience Squared (units of 100)	-0.0208	-0.0206	-0.0044*	-0.0046*
	[0.0147]	[0.0149]	[0.0025]	[0.0025]
2007	0.0122*	0.0123*	0.0131**	0.0132**
	[0.0065]	[0.0065]	[0.0055]	[0.0055]
2008	0.0126*	0.0126*	0.0159**	0.0162**
	[0.0075]	[0.0075]	[0.0063]	[0.0063]
Female	0.0147**	0.0144**	0.0096**	0.0086*
	[0.0061]	[0.0061]	[0.0049]	[0.0049]
Constant	-0.0202**	-0.0225**	-0.0148*	-0.0128
	[0.0101]	[0.0104]	[0.0078]	[0.0080]
Observations	5,916	5,916	13,407	13,407
R-squared	0.008	0.007	0.009	0.007

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 18: Lit testing tails, school-grade level results with clustered error terms

Variables	Praxis I-10%	Praxis I-25%	Praxis II-10%	Praxis II-25%
Alternative Certification	-0.0170**	-0.0166**	-0.0134*	-0.0135*
	[0.0076]	[0.0076]	[0.0076]	[0.0076]
Alt. Cert/Exp. Interaction Term	0.0054**	0.0054**	0.0042*	0.0045**
	[0.0022]	[0.0022]	[0.0022]	[0.0021]
Licensure Test-Top	0.0084	0.0104**	0.0086*	0.0043
	[0.0067]	[0.0047]	[0.0045]	[0.0033]
Licensure Test-Bottom	-0.0158***	0.0022	-0.0138***	-0.0084**
	[0.0058]	[0.0048]	[0.0046]	[0.0034]
BA Not earned in Arkansas	0.0080	0.0080	-0.0018	-0.0018
	[0.0099]	[0.0100]	[0.0044]	[0.0044]
Graduate Degree	-0.0040	-0.0040	0.0004	0.0003
	[0.0042]	[0.0042]	[0.0031]	[0.0031]
Experience	0.0007	0.0006	0.0006	0.0006
	[0.0012]	[0.0012]	[0.0005]	[0.0005]
Experience Squared (units of 100)	-0.0003	-0.0003	-0.0007	-0.0008
	[0.0084]	[0.0084]	[0.0018]	[0.0018]
2007	0.0012	0.0014	0.0020	0.0021
	[0.0043]	[0.0043]	[0.0038]	[0.0038]
2008	-0.0007	-0.0005	-0.0021	-0.0020
	[0.0053]	[0.0053]	[0.0045]	[0.0045]
Female	0.0090**	0.0085*	0.0048	0.0042
	[0.0045]	[0.0045]	[0.0036]	[0.0036]
Constant	-0.0082	-0.0112	-0.0056	-0.0044
	[0.0071]	[0.0072]	[0.0057]	[0.0058]
Observations	5,916	5,916	13,407	13,407
R-squared	0.007	0.006	0.004	0.004

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

High School

High school students take state standardized tests in four areas: algebra, biology, geometry, and literacy. The algebra, biology, and geometry exams are end-of-course exams, thus students may take them in different grades, depending on when they take the course. In Arkansas, Algebra is typically taken in 9th grade, but a significant number of students do take the

class in 8th grade. Geometry is typically taken in 10th grade. The literacy exam is taken by all 11th grade students. I do not examine teacher impacts in biology.

In this section, I present the analysis of end-of-course exams in Algebra and geometry, and the eleventh grade literacy exam separately. In each of these analyses, I use most of the same variables that I used in the Benchmark analyses: licensure route, graduate degrees, experience, Praxis I, and Praxis II professional knowledge. Additionally, I am able to utilize the Praxis II content scores for each subject. For a small sample of secondary teachers, I am also able to identify the college major by a Classification of Instructional Program (CIP) code. These codes indicate if a teacher earned their degree in an education field or in the content field. For these analyses, I group all math content majors together and all English language arts majors together. I am also able to identify if the secondary teachers earned an education degree in their content field. A teacher who earns a bachelor of science in math education, for example, would fall in this category.

Algebra

We would expect teachers of Algebra to have good math skills, at least in comparison to teachers of other subjects. A good measure of this is the Praxis I math exam. As it turns out, the Algebra teachers do score markedly higher than other teachers. In Table 19, I display descriptive statistics of the teachers in the least restrictive Algebra model. As the data show, traditionally certified algebra teachers scored 0.899 standard deviations above the mean on the Praxis I math test, while alternatively certified algebra teachers scored 1.213 standard deviations above the mean of all test teachers in the data set. That is, Algebra teachers from both routes score much higher on the Praxis I exam than teachers of other subjects.

On each of the licensure exams, I conducted T-tests to determine if the differences between traditionally and alternatively certified teachers were significant. On all three Praxis I tests of basic skills, alternatively certified teachers scored significantly higher than traditionally certified teachers, *p.* <0.01. There was not a statistically significant difference between the two groups on the Praxis II professional knowledge exam, but alternatively certified teachers scored higher on the math content exam, a difference that was marginally significant, *p.*<0.10.

The two groups differed from each other in other notable ways. Just as they did at the elementary level, traditionally certified algebra teachers have significantly more experience. In the same vein, traditionally certified algebra teachers are more likely to have a graduate degree, which would make sense if they have had more years to complete an advanced degree.

Table 19: Characteristics of Algebra Teachers in Least Restrictive Model

Teacher Characteristic	Traditional	Non-Traditional	Total
Algebra Teachers	3,155	235	3,390
Teachers with a Graduate Degree	41.2%***	30.6%	40.5%
Avg. Praxis I-Math z-score	0.899 (0.821)***	1.213 (0.566)	0.043 (0.797)
Avg. Praxis I- Writing z-score	0.199 (1.001)***	0.443 (0.818)	0.234 (0.980)
Avg. Praxis I- Reading z-score	0.197 (0.824)***	0.591 (0.549)	0.252 (0.803)
Avg. Praxis II- Prof. Know. z-score	0.062 (0.749)	0.114 (0.754)	0.064 (0.749)
Avg. Praxis II- Math	-0.061 (0.723)*	0.036 (0.565)	-0.053 (0.712)
White	89.1%*	92.8%	89.4%
Black	9.2%***	2.6%	8.8%
Female	65.9%	60.9%	65.6%
Years of Experience	13.3***	2.7	12.5

*** p<0.01, ** p<0.05, * p<0.1

The difference in terms of effectiveness between alternatively certified teachers of algebra and traditionally certified teachers is negligible. Indeed, alternatively certified teachers appear to be no more or less effective than traditionally certified teachers as the coefficients for alternatively certified teachers are mixed and none are significant. Likewise, there seems to be no difference in terms of effectiveness for teachers with a graduate degree.

Just as they were in the Benchmark analyses, a teacher's prior test scores on licensure exams tend to be the strongest predictors of effectiveness in the classroom. Although, in this instance it is not the Praxis II professional knowledge test that is positively correlated with value-added. Rather, Praxis II math content tests and an average of Praxis I scores are positively associated with student achievement. When each of these exams is included as the sole test score, both are significant at conventional levels. When all three exams are included in the analysis, Praxis I and Praxis II content knowledge tests are marginally significant. Interestingly, the relationship between teachers' scores on the professional knowledge exam and value-added student achievement is negatively related and marginally significant when the other two exams are included in model five.

Table 20: Algebra Results with error terms clustered at the school-grade level

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Alternative Certification	-0.008	-0.017	0.011	-0.013	-0.026	0.083
	[0.021]	[0.035]	[0.071]	[0.035]	[0.036]	[0.086]
Alt. Cert/Exp. Interaction Term		0.002	-0.009	-0.009	0.003	-0.025
		[0.009]	[0.014]	[0.007]	[0.010]	[0.016]
Praxis II- Math					0.024**	0.033*
					[0.009]	[0.017]
Praxis II- Prof. Know.			0.006			-0.035*
			[0.009]			[0.020]
Praxis I-Average				0.033***		0.039*
				[0.013]		[0.021]
BA Not earned in AR		-0.006	0.005	-0.063	-0.012	-0.067
		[0.016]	[0.018]	[0.040]	[0.020]	[0.045]
Graduate Degree		0.003	0.005	0.010	-0.001	-0.005
		[0.012]	[0.013]	[0.017]	[0.013]	[0.021]
Experience		0.002	0.003	0.013***	0.002	0.008
		[0.002]	[0.002]	[0.005]	[0.002]	[0.007]
Experience Squared (units of 100)		-0.006	-0.018***	-0.105***	-0.006	-0.080**
		[0.005]	[0.007]	[0.035]	[0.007]	[0.040]
2007		0.006	0.000	0.028	0.000	0.033
		[0.013]	[0.015]	[0.018]	[0.013]	[0.020]
2008		-0.007	-0.015	0.016	-0.010	0.022
		[0.015]	[0.017]	[0.022]	[0.015]	[0.025]
2009		0.012	0.006	0.027	-0.004	0.030
		[0.016]	[0.017]	[0.023]	[0.016]	[0.027]
2010		-0.000	0.006	0.017	-0.003	0.037
		[0.020]	[0.021]	[0.029]	[0.019]	[0.034]
Female		0.043***	0.047***	0.021	0.044***	0.029
		[0.012]	[0.015]	[0.019]	[0.014]	[0.024]
Constant	-0.014	-0.039*	-0.036	-0.067**	-0.022	-0.041
		[0.021]	[0.023]	[0.030]	[0.022]	[0.039]
Observations	4,012	3,390	2,070	1,381	2,636	832
R-squared	0.001	0.012	0.024	0.030	0.017	0.054

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Additional analyses were conducted while restricting the sample to only teachers in their first five years of teaching. In these analyses, none of the variables of interest were significant and none of the signs changed dramatically.

Geometry

Like Algebra teachers, we might expect teachers of geometry to have higher scores on math exams than the average teacher. Once again, this bears out in the data, both traditionally and alternatively certified geometry teachers score relatively high on the Praxis I math exam, 0.928 of a standard deviation for traditionally certified teachers and 1.261 for alternatively certified teachers. In fact, both sets of teachers score higher than the average teacher on all three segments of the Praxis I exam; though in each instance, alternatively certified teachers scored significantly higher than the traditionally certified teachers. On the Praxis II professional knowledge exam, the two groups were closer to the average and the difference between the two groups was not significant, nor were they significantly different on the Praxis II math content knowledge exam. Alternatively certified teachers were significantly more likely to be non-black, male, and have significantly less experience.

Table 21: Characteristics of Geometry Teachers in Least Restrictive Model

Teacher Characteristic	Traditional	Non-Traditional	Total
Geometry Teachers	2,437	204	2,641
Teachers with a Graduate Degree	43.0%	30.9%	42.0%
Avg. Praxis I-Math z-score	0.928 (0.765)***	1.261 (0.522)	0.982 (0.741)
Avg. Praxis I- Writing z-score	0.137 (0.888)***	0.460 (0.897)	0.189 (0.897)
Avg. Praxis I- Reading z-score	0.239 (0.788)***	0.507 (0.536)	0.282 (0.759)
Avg. Praxis II- Prof. Know. z-score	0.087 (0.694)	0.198 (0.752)	0.091 (0.697)
Avg. Praxis II- Math	-0.035 (0.700)	0.017 (0.559)	-0.030 (0.688)
White	89.4%*	93.6%	89.7%
Black	9.2%***	0.5%	8.5%
Female	65.1%***	51.0%	64.0%
Years of Experience	13.6***	3.0	12.8

We might expect the geometry results to be very similar to the algebra results. After all, both are end-of-course exams in math subjects and the samples of teachers in both subjects are quite similar. Nevertheless, there are some notable differences between the algebra and geometry

results. For starters, the coefficient on alternatively certified teachers is negative in every model and marginally significant in two of those models, including the most direct test of certification. Unlike the algebra results, none of the Praxis scores are significant predictors of effectiveness in teaching geometry. Similar to the other end-of-course analysis, teachers of geometry tend to benefit from years of experience, but are not more effective when they have a master's degree.

Table 22: Geometry Results with error terms clustered at the school-grade level

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Alternative	-0.040*	-0.032	-0.093*	-0.001	-0.041	-0.104
Certification	[0.021]	[0.036]	[0.053]	[0.038]	[0.039]	[0.073]
Alt. Cert/Exp.		-0.005	0.014	-0.014	-0.002	0.014
Interaction Term		[0.011]	[0.009]	[0.011]	[0.012]	[0.011]
Praxis II- Math					0.014	0.011
					[0.010]	[0.020]
Praxis II- Prof.			0.018			-0.003
Knowledge			[0.012]			[0.025]
Praxis I-Average				0.017		0.041*
				[0.018]		[0.024]
BA Not earned in		-0.037**	-0.031	-0.088*	-0.047**	-0.032
Arkansas		[0.018]	[0.025]	[0.051]	[0.022]	[0.070]
Graduate Degree		0.0073	0.002	0.020	-0.006	0.013
		[0.012]	[0.014]	[0.020]	[0.014]	[0.026]
Experience		0.004**	0.004	0.017***	0.004*	0.013
		[0.002]	[0.003]	[0.005]	[0.002]	[0.008]
Experience Squared		-0.014**	-0.018**	-0.106***	-0.011	-0.090**
(units of 100)		[0.005]	[0.008]	[0.037]	[0.007]	[0.011]
2007		0.014	-0.005	0.005	0.006	-0.031
		[0.018]	[0.021]	[0.027]	[0.019]	[0.029]
2008		0.013	0.008	0.014	0.005	0.002
		[0.024]	[0.027]	[0.030]	[0.023]	[0.030]
2009		0.002	-0.029	0.003	-0.014	-0.035
		[0.022]	[0.027]	[0.031]	[0.021]	[0.030]
2010		0.006	-0.019	0.013	-0.001	-0.011
		[0.023]	[0.027]	[0.032]	[0.022]	[0.031]
Female		0.018	0.025	0.018	0.020	0.014
		[0.012]	[0.015]	[0.018]	[0.013]	[0.022]
Constant	-0.004	-0.024	0.005	-0.067*	-0.011	-0.014
	[0.011]	[0.026]	[0.032]	[0.036]	[0.025]	[0.044]
Observations	3,091	2,641	1,552	1,071	2,109	652
R-squared	0.003	0.019	0.025	0.049	0.023	0.053

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Once again, I conducted the above analyses while restricting the sample to novice teachers in their first five years. As it was in algebra, none of the variables of interest were significant at conventional levels.

English Language Arts

Unlike algebra and geometry, the English language arts exam is not an end-of-course exam; rather, it is an exam given to all students in the eleventh grade. Though this is a notable difference, I do not expect this to markedly change the results. A more important distinction between the end-of-course exams and the English language arts exam is in regards to the teaching core. The number of alternatively certified teachers is markedly lower in ELA than in the two math subjects. The likely reason for this is simple supply and demand. The Arkansas Department of Education regularly reports shortages of secondary math teachers, but rarely reports shortages in English subjects (Arkansas Department of Education 2012). More shortages in math would necessitate the hiring of more alternatively certified teachers.

In both math subjects, alternatively certified teachers scored significantly higher on all sections of the Praxis I; not so in ELA. The average score was higher for non-traditional teachers than it was for traditionally certified teachers; but the difference between the two was not significant. Alternatively certified teachers did score significantly higher, however, on both types of Praxis II exams, professional knowledge and the content exam. For the professional knowledge exam, the average score for alternatively certified teachers was 0.644 standard deviations higher than for traditionally certified teachers.

Table 23: Characteristics of ELA Teachers in Least Restrictive Model

Teacher Characteristic	Traditional	Non-Traditional	Total
English 11 Teachers	1,205	76	1,281
Teachers with a Graduate Degree	41.5%	35.5%	41.1%
Avg. Praxis I-Math z-score	0.188 (0.883)	0.311 (0.854)	0.202 (0.879)
Avg. Praxis I- Writing z-score	0.486 (0.886)	0.5499 (0.612)	0.493 (0.859)
Avg. Praxis I- Reading z-score	0.453 (0.717)	0.584 (0.642)	0.468 (0.709)
Avg. Praxis II- Prof. Know. z-score	0.180 (0.798)***	0.824 (0.607)	0.204 (0.801)
Avg. Praxis II- ELA	-0.043 (0.757)***	0.249 (0.747)	-0.021 (0.759)
White	90.0%	89.5%	89.9%
Black	8.7%**	1.3%	8.3%
Female	85.6%***	71.1%	84.8%
Years of Experience	12.6***	3.3	12.1

The most notable finding in terms of ELA is that none of the observed teacher characteristics appear to significantly influence student achievement. That is, none of the variables are significant, save for the binary variable that indicates if teachers earned their bachelor's degree outside of Arkansas. Even then, the coefficient is only marginally significant in two of the five models. Performance on the Praxis I and II exams is not significant, nor are graduate degrees or experience. The results were similar when the data were restricted to novice teachers.

Table 24: ELA Results with error terms clustered at the school-grade level

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Alternative Certification	0.006	0.014	-0.069	0.010	0.008	-0.176
	[0.17]	[0.042]	[0.099]	[0.051]	[0.043]	[0.136]
Alt. Cert/Exp.		0.002	0.018	-0.001	0.002	0.040
Interaction Term		[0.009]	[0.021]	[0.013]	[0.010]	[0.028]
Praxis II- English					0.013	
Language Arts					[0.009]	
Praxis II- Professional			0.015			0.013
Knowledge			[0.011]			[0.019]
Praxis I-Average				0.005		-0.030
				[0.013]		[0.031]
BA Not earned in		-0.024*	-0.032*	0.004	-0.024	-0.044
Arkansas		[0.015]	[0.017]	[0.025]	[0.016]	[0.036]
Graduate Degree		-0.005	0.014	-0.010	0.003	-0.011
		[0.008]	[0.010]	[0.015]	[0.010]	[0.021]
Experience		0.001	0.002	-0.001	0.001	0.001
		[0.001]	[0.002]	[0.004]	[0.002]	[0.006]
Experience Squared		-0.001	-0.007	0.005	-0.002	-0.006
(units of 100)		[0.004]	[0.007]	[0.020]	[0.005]	[0.024]
2008		0.040	0.106	0.042	0.009	0.060
		[0.111]	[0.116]	[0.141]	[0.105]	[0.176]
2009		0.039	0.108	0.046	0.013	0.070
		[0.111]	[0.117]	[0.142]	[0.106]	[0.178]
2010		0.043	0.110	0.042	0.011	0.060
		[0.110]	[0.116]	[0.141]	[0.105]	[0.179]
Female		0.019	0.019	0.022	0.014	0.067
		[0.015]	[0.018]	[0.025]	[0.018]	[0.046]
Constant	0.002	-0.052	-0.135	-0.055	-0.023	-0.110
	[0.009]	[0.111]	[0.120]	[0.144]	[0.106]	[0.181]
Observations	1,539	1,281	739	530	991	293
R-squared	0.000	0.011	0.042	0.008	0.010	0.049

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

College Major

It makes intuitive sense that a teacher with a great deal of content knowledge would be more effective than a teacher with little content knowledge. In the following analyses, I test this assumption. As I have noted, I am able to identify the degree of a set of teachers via a CIP code. This code indicates whether the teacher has an education degree or a content area degree. In the

analyses below, I examine the impact of having a math degree from the content area or a math education degree on value-added student achievement. For the 11th grade literacy exam, I conduct the same type of analyses with English language arts degrees. Table 25 indicates the percentages of teachers with a degree in their content field or in an education content field. A content degree would mean a math or language arts degree. An education content degree is an education degree in a specific field. The remaining teachers have an assortment of other degrees.

Table 25: Number of teachers with a content or education content degree in teaching area

	Algebra	Geometry	ELA
Content Degree	13.8%	12.4%	18.6%
Ed. Content Degree	16.2%	18.2%	22.7%

Table 26: Impact of College Degrees on Value-added Student Achievement, with error terms clustered at the school-grade level

Variables	Algebra	Geometry	ELA
Content Degree	0.064**	0.002	0.043
	[0.032]	[0.042]	[0.035]
Ed. Content Degree	-0.022	-0.035	0.017
	[0.035]	[0.040]	[0.027]
Experience	0.015**	0.025**	0.001
	[0.007]	[0.011]	[0.006]
Experience Squared	-0.064	-0.164**	0.020
(units of 100)	[0.046]	[0.065]	[0.036]
2007	0.024	-0.007	
	[0.040]	[0.063]	
2008	-0.038	-0.042	-0.010
	[0.039]	[0.054]	[0.079]
2009	-0.004	-0.039	-0.024
	[0.040]	[0.052]	[0.078]
2010	-0.042	-0.040	-0.021
	[0.041]	[0.057]	[0.078]
Constant	-0.025	0.017	-0.000
	[0.039]	[0.056]	[0.080]
Observations	419	346	172
R-squared	0.050	0.073	0.041

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

The analyses indicate that teachers with a degree in a math content area tend to perform significantly better than the average teacher and better than teachers with a math education degree in Algebra. This is in line with previous research. In three separate studies, Goldhaber and Brewer (1997a; 1997b; 2000) found that secondary teachers with a degree in math were significantly more effective than those without a math degree. Harris and Sass (2011) similarly found that math teachers with more college credits in math tended to perform better than teachers with fewer credits. Conversely, Goldhaber and Brewer (2000) and Rockoff et al. (2008) found no benefit to having a degree in education.

Prior research has tended to find no impact for teachers of English language arts with a degree in the field. Likewise, I find no impact of having an ELA content degree or an ELA education degree. It should be noted, however, that the number of observations in each of the three observations is below 420 teachers, only 172 in ELA.

Summary of Findings

Ensuring students have a quality teacher is a worthy endeavor. To do this, Arkansas, like most states, has instituted a licensure system whereby teachers must pass a series of examinations to earn a teaching license. The traditional route requires a degree in education, student teaching, and passage of the licensure tests. The alternative route allows individuals with any degree, who pass the licensure exams, and find a job, to enter the classroom without student teaching experience.

In this chapter I address my research questions and sub questions. To answer these questions I utilize a two-step approach. First, I estimate student value-added with a parsimonious model, controlling for two years of achievement data in math and ELA. I attribute the value-

added to teachers at a school-grade level and regress on observable teacher characteristics, including licensure exam scores and licensure route. A summary of the findings can be seen in Table 27. In the table, I note variables that were significant and positive with a "+", while variables that were significant and negative are marked with a "-". If a variable was not significant, the space is left blank. For the Benchmark exams, I did not analyze the effect of Praxis II content scores, thus I have marked it NA in each of those models.

Table 27: Summary of Findings

	Praxis I	Alt Cert	Praxis II-Prof.	Praxis II-Content	Graduate Deg	Exp
Math (School-Grade)	+		+	NA		
ELA (School-Grade)	+		+	NA		
Math (S-G-Subject)	+		+	NA		
ELA (S-G-Subject)						
Math (Novice)	+		+	NA		
ELA (Novice)	+		+	NA		
Algebra	+			+		+
Geometry		-				+
11th Grade Literacy						

Below I answer the research questions outlined in chapter one.

(1). Do traditionally certified teachers generate significantly larger increases in student achievement on standardized exams than alternatively certified teachers?

Certification route does not provide clear results. The differences between groups tend to be small and marginally significant only when I control for prior academic achievement as measured by teacher licensure exams. Since alternatively certified teachers score significantly higher on licensure exams, on average, including these scores biases the estimates of alternative certification downward. I will discuss this further below. Nevertheless, the coefficient on alternative certification typically remains negative, but insignificant, when teacher test scores are not included. At the secondary level, the differences between traditionally and alternatively

certified teachers are not significantly different at the *p.* <0.05 level, even when controlling for teacher performance on licensure exams. In geometry, alternatively certified teachers performed less well, and the difference was marginally significant at the *p.* <0.10 level. Therefore, I fail to reject the null hypothesis that alternatively certified teachers increase student achievement at similar rates as traditionally certified teachers in 8 out of 9 tests of that hypothesis. On average, alternatively certified teachers tend to perform slightly lower than traditionally certified teachers, but there is more variation within each group than between groups.

(2). Are teacher licensure exams scores related to gains in student achievement?

Yes, the data are sufficient to reject the null hypothesis that there is no relationship between a teacher's score on a licensure exam and their effectiveness at improving student achievement, although there are some caveats to this conclusion. Teachers who score higher on licensure exam tend to perform better than their lower scoring counterparts. The Praxis I test has a positive significant relationship with student achievement growth in math and ELA on the Benchmark exam. At the secondary level, the Praxis 1 has a positive significant relationship with student growth on the algebra EOC, but no on the geometry EOC or 11th grade literacy exam. At the elementary level, the Praxis II professional knowledge exam is a significant predictor of performance in teaching both math and ELA. At the secondary level, it is not significantly related to student achievement growth. I am only able to analyze the Praxis II content exam at the secondary level. Here, it is significantly related to student achievement growth in algebra, but not in geometry or 11th grade literacy.

It is important to remember, licensure exams were not intended to be predictors of performance in the classroom; rather they are a minimum quality guard. Thus, they weed out extremely low performers. Additionally, since they are designed around a cut-point, some tests

have ceiling effects. Still, performance on licensure exams, especially the Praxis II, tends to be a significant predictor of future performance. This is in accord with a wide literature base that suggests teachers with more content knowledge or higher prior academic achievement tend to be more effective teachers. I suspect a test designed to capture a full range of abilities among prospective teachers may have even better predictive power.

(3). Do factors that determine teacher salaries impact student achievement?

Although the findings in this dissertation are not as conclusive as some of the previous research, there is some reason to believe experience matters, at least in the first few years of teaching. Having a graduate degree, on the other hand, does not seem to improve a teacher's performance in the classroom.

(4). Do secondary teachers with a degree in the field they are teaching generate the same learning gains as teachers without a degree in the content field?

Algebra teachers with a degree in their content field perform significantly better than algebra teachers without a degree in the content field. However, in geometry and 11^{th} grade literacy, no such relationship was detected. Interestingly, teachers who earn a degree from a college of education with an emphasis in their content field did not perform significantly better or worse than other teachers in the sample.

Alternative Certification and Licensure Exam Scores

One of the promises of alternative certification is the ability to bring high-performing individuals from other fields into teaching, without requiring them to get an education degree. From our analyses, this seems to be happening. In every analysis, the individuals entering the profession via an alternative route score higher on at least one of the licensure exams. This is not out of the ordinary. In New York, "only 5 percent of newly hired Teaching Fellows and TFA

teachers in 2003 failed the LAST [Liberal Arts and Sciences Test] exam on their first attempt, while 16.2 percent of newly hired traditional teachers failed the LAST exam" (Boyd et al., 2008, p. 815).

At the outset of the dissertation, I indicated the analysis of licensure routes could not account for the selection effects of each route to the classroom. That means a comparison of teachers entering the classroom via the traditional or alternative route does not just capture the effect of the training, or lack thereof, which is entailed in each pathway. The comparison does capture those differences, but it also captures the differences that are caused by systematic selection into the two fields. As we have seen, alternatively certified teachers consistently score higher, on average, on most licensure exams. This is a very good indicator that a very different type of individual is entering the classroom through the alternative pathways.

For this reason, it is inappropriate to control for licensure exam score when comparing the two pathways. Alternative certification attracts individuals with higher scores on licensure exams and performance on these exams is correlated with performance in the classroom. Thus, controlling for licensure scores in this case would be like controlling for batting average in a regression model predicting the number of victories achieved by a baseball team.

Another systematic difference between teachers entering through traditional and alternative pathways is gender. The alternatively certified teachers in this sample are significantly more likely to be male, especially at the elementary level. Thus, the alternative licensure route seems to be attracting above average individuals, in terms of licensure exam scores, and males into elementary classrooms.

Chapter V - Discussion

It is a well-established fact that teachers can have a tremendous impact on student achievement. In this dissertation, I examine the relationship between several observable teacher characteristics and value-added student achievement. The characteristics I examine are common metrics used to identify teacher quality: certification, licensure exam scores, experience, master's degrees, and content area degrees. Certification and licensure exams scores are screens, which are used to keep low-performing individuals from entering the profession. Experience and master's degrees are used to determine the salary of the vast majority of teachers (National Center for Education Statistics 2008). Since they are used to determine salaries, they are *de facto* measures of quality. Content degrees are not used to determine eligibility for certification in Arkansas, nor are they used for pay decisions. They are, however, often lauded as desirable, especially for secondary teachers.

In this chapter, I briefly summarize the findings from this dissertation and describe how those findings fit within the larger body of research. From these findings, I outline some policy suggestions for Arkansas lawmakers. Based on the evidence, these policy recommendations would improve the overall quality of teachers, a desirable goal. I recognize, however, that this study has numerous limitations which warrant caution in both the interpretation of my findings and in the policy recommendations. I believe, however, that with additional study, these findings will hold. I offer the following recommendations for future study. These recommendations are not just for scholars, but also for state officials who oversee and maintain the collection of data. I close the chapter with some final thoughts.

Teacher Certification

In this dissertation, I sought to find out if any observable teacher characteristics were significantly related to teacher effectiveness. To clarify, by effectiveness I mean the ability of a teacher to improve student achievement on state administered standardized achievement exams. While there may be many definitions of effectiveness, I believe improving student achievement is the most important. As Ritter and Shuls (2012) noted, "If a teacher teaches what seems to him a very thoughtful and careful lesson, but students don't learn the skills and/or knowledge intended, was it a good lesson? The answer, we believe, is obvious: If students aren't learning, then the teaching isn't effective" (p. 34). Ritter and Shuls go on to make a compelling case for why test scores are an appropriate measure of teacher effectiveness. They note that all other strategies for capturing effectiveness, like observations or National Board Certification, are all proxies intended to capture whether or not a teacher is an effective educator. Obviously, I concur with Ritter and Shuls. The tests used in Arkansas align to our curriculum standards. Teachers are hired to teach those standards; therefore, evaluating their performance based on student growth seems prudent.

The primary focus of this dissertation is on the relationship between licensure screens and effectiveness. I examine whether traditionally certified teachers are more effective than alternatively certified teachers. I also examine the relationship between a teacher's performance on licensure exams and student achievement. The findings presented here are interesting in and of themselves, but also because they explore whether the findings from the more urbanized states and cities are consistent with those of an analysis conducted in Arkansas, a rural state.

Results

Based on the evidence presented in this dissertation, the difference between traditionally and alternatively certified teachers in terms of effectiveness is negligible. I examine the difference between these two groups in mathematics and language arts via the Benchmark exams administered in grades three through eight, on end-of-course exams in Algebra and geometry, and on the grade 11 literacy exam. In every instance, alternatively certified teachers were not significantly better or worse than traditionally certified teachers at the standard significance level of $p. <0.05$. There was one marginally significant difference in geometry, where traditionally certified teachers tended to perform at higher levels.

In the broader research literature, the results have not been conclusive on this matter. Some researchers have found alternatively certified teachers to be significantly worse (Boyd et al. 2008; Clotfelter, Ladd, and Vigdor 2007 & 2010; Goldhaber & Anthony 2007). Others have null findings, as I have (Croninger et al. 2003; Goldhaber & Brewer 2000; Constantine et al. 2009). All of these studies, even the ones with significant findings, note that the difference within groups is much larger than the difference between groups. I note this same phenomenon here. In short, my findings demonstrate the differences between traditionally and alternatively certified teachers are negligible and small in comparison to the variation within each group. These findings fit within the existing research base.

Teacher Licensure Exams

Licensure exams are designed to keep low-performing individuals from the classroom; they are not designed to be measures of teaching effectiveness. As I have noted, the structure of these exams may mean that some exams have ceiling effects, because the test is not attempting to

determine exactly how much someone knows. Nevertheless, licensure exams do provide useful information regarding teacher quality.

In this dissertation, I analyze the relationship between a teacher's performance on licensure exams and value-added student achievement. In grades three through eight, I utilize teacher data from the Praxis I exam and the Praxis II professional knowledge exam. Praxis I is a test of basic skills in math, reading, and writing. Teachers generally take this in their sophomore year to enter into a teacher preparation program. The Praxis II professional knowledge exam is typically taken at the end of an undergraduate program. It is meant to assess the teacher's knowledge of pedagogical practice. In addition to these two exams, I examine the relationship between the Praxis II content exams at the high school level.

Results

Although it is designed as a test of basic skills, the Praxis I exam was consistently related to value-added student achievement in a significant positive direction. It is significantly related in both math and language arts at the elementary level. Praxis I scores are also significantly related to effectiveness in teaching Algebra at the secondary level. The Praxis II professional knowledge exam, on the other hand, is significantly related to a teacher's performance in the elementary grades, but not secondary. Finally, Algebra teachers who scored higher on the Praxis II math content exam tend to perform better. This finding does not hold for geometry teachers or for English language arts teachers.

These findings are not altogether surprising. The Praxis I exam is a skills-based test, meaning it is designed to assess what a teacher is able to do or how well they can perform in each of the three subjects. That is not too unlike the SAT (Boyd et al 2008) or the ACT (Ferguson & Ladd 1996), which have been found to be positively related to teacher

effectiveness. Individuals who are able to perform better on a standardized test of this sort may be more academically capable than students who do not score so well. It makes sense that more academically capable people would be more effective instructors.

The Praxis II results are also not surprising. Numerous researchers have found a significant, positive relationship between performance on a teacher licensure exam and a teacher's ability to improve student achievement at the elementary level (Clotfelter, Ladd, & Vigdor 2006, 2007; Goldhaber 2007) and high school levels (Clotfelter, Ladd, & Vigdor 2010). The findings also fit within the theoretical understanding of what it might take to be an effective elementary or secondary teacher. Shuls and Ritter (2013) state, it is evident "that the prerequisites to be an effective elementary or secondary teacher are markedly different" (p. 32). We suggest that elementary teachers need a deeper understanding of child development than high school teachers. Conversely, we argue that high school teachers need to have a good grasp of content knowledge. The results presented in this dissertation support those assertions. Elementary teachers who score higher on the Praxis II professional knowledge exam tend to be more effective. This does not carry through to high school teachers. I am unable to examine the Praxis II content knowledge exam at the elementary grades, but at the secondary there is a significant relationship between a teacher's performance on the math content knowledge exam and their effectiveness at teaching algebra.

Experience and Masters Degrees

The Schools and Staffing Survey (2008), released by the National Center for Education Statistics, indicates that the vast majority of teachers are paid based on a salary schedule. Though these schedules can take many forms, the most common form includes steps and lanes based on years of experience and additional college attainment, including a master's degree. Because

many school districts base almost their entire teacher compensation on these two variables, they are by proxy traditional measures of teacher quality.

Results

There is some evidence, from my analyses, that experience matters to a limited extent. The literature is quite clear that new teachers tend to perform worse, on average, than their more experienced counterparts (Boyd et al., 2006; Boyd et al. 2008; Clotfelter, Ladd, & Vigdor 2006, 2007, Croninger et al. 2003; Rockoff 2004; Rivkin, Hanushek, & Kain 2005; Kane, Rockoff, & Staiger 2008; Krieg 2006; Goldhaber & Anthony, 2007; Cavalluzzo, 2004; West & Chingos, 2009; Dee, 2004; Xu, Hannaway, & Taylor, 2011). The benefit to a year of experience, however, is not linear; most of the gains come within the first few years of teaching.

Master's degrees are not significantly related to student achievement in any of my models. This too is consistent with the literature (Aaronson, Barrow, & Sander 2007; Clotfelter, Ladd, & Vigdor 2007, 2010; Croninger et al. 2003; Hanushek et al. 2005; Krieg 2006; Rockoff et al. 2008; Xu, Hannaway, & Taylor, 2011). There are many reasons why this may be the case. First, it could be that teachers take master's courses to earn a raise, not to improve. That is an important distinction, as any teacher will tell you, "you get out of class, what you put into it." If the goal is not improvement, we might not expect dramatic improvements in effectiveness. Moreover, teachers often earn advanced degrees in subjects that may not be designed to improve their instructional capacity. For instance, many teachers earn master's degrees in administration. It may be the case that the training necessary to become an administrator is not related to the skills necessary to be a good teacher.

Content Degrees

Intuitively, it makes sense that effective teachers will have a deep understanding of content knowledge. The National Math Panel concluded that it is not just intuition, but evidence that supports this claim (U.S. Department of Education 2008). The panel suggests that teachers of math need to have a deep understanding of math. The Arkansas data provide the ability to test this assumption with a sample of secondary teachers for whom the appropriate data are available. Using CIP codes I am able to identify whether a teacher has earned a bachelor's degree in a content field or in the field of education. Moreover, I am able to identify whether the teachers with an education degree have earned an education degree with a focus in their content area; such as a bachelor's of science in math education.

Results

For the algebra and geometry end-of-course exams, I examine the relationship between having a bachelor's degree in math or math education and effectiveness at improving student achievement. The relationship between a content degree and value-added student achievement is positive and statistically significant in Algebra, $.p < 0.05$. It is positive in geometry, but not significant. Interestingly, the coefficient for teachers with a math education degree is negative, but not significant, in both algebra and geometry. Similarly, there is not a significant relationship between language arts degrees and performance as an 11^{th} grade language arts teacher.

It is not surprising that there is a significant result in math. Goldhaber and Brewer (1997a; 1997b; 200) similarly conclude that math educators tend to be more effective if they have a degree in the field. Likewise, Harris and Sass (2011) note that teacher effectiveness is positively related to the number of subject content credits a teacher has completed. No such findings have been demonstrated in language arts.

Limitations and Suggestions for Future Research

There are, of course, limitations to these findings that warrant caution. One big limitation is the inability to match students directly to teachers. Although matching teachers and students to a school grade does not bias the results, it makes the difference between groups more difficult to identify. It also makes it difficult to generate precise estimates of the impact of a particular variable. Future research should attempt to match students to teachers. To do this, the Arkansas Department of Education will need to improve data collection and verify the links between students and teachers are accurate.

The location of this study also provides a limitation on interpretation. Just as New York City is not like much of the country, Arkansas does not represent urban states and cities. Still, the results here in many ways suggest the findings are consistent between urban and rural settings. Additional research in other states will help verify these findings.

There are also limitations that arise from the nature of the data. For instance, I have a binary variable that indicates if a teacher has an alternative or traditional teaching license. From this we see almost no real differences between the groups. What we do not see is the effectiveness of individuals with no certification. It could be the case that both pathways do an effective job of keeping out the lowest performers. I also have the problem of non-random sorting of teachers to licensure paths and students to teachers. I cannot say that one pathway is more effective in training teachers because the individuals that select into each path are noticeably different. Therefore the evaluation has to take into account the differences of the individuals selecting into each licensure route. The issue of non-random sorting of students to teachers can, and has been (Constantine et al. 2009) addressed by conducting an experiment

where students are randomly assigned to teachers of different routes. To date, non-random sorting at the teacher level has not been addressed.

The findings on teacher licensure exams are also limited by the data. Praxis exams are not intended to capture an individual's full range of ability. In many cases these tests have ceilings, with many individuals scoring towards the highest score possible. Teacher licensure scores are also limited in that teachers who fail the exams do not appear in my data. Since a pattern exists in the existing data between scores and performance in the classroom, I expect that pattern to continue in both directions, but I cannot know that.

Finally, the data are limited for individuals with master's degrees and content areas. Once again, the Arkansas Department of Education could help improve our knowledge in this area by keeping better track of the degree that a teacher earns. Though master's degrees do not seem to help teachers improve, I cannot say whether a master's degree in math would help a math teacher or the effect of other specific degrees.

Policy recommendations

Although there are limitations to the results discussed here, these findings are consistent with the existing research. Together, the findings presented in this dissertation and the literature make it clear that changes need to be made to current policy. In this section, I offer some policy solutions that I believe will help improve the quality of the teacher workforce.

Expand Pathways to the Classroom

In the beginning of this dissertation, I note there are two distinctive camps when it comes to improving teacher quality. Shuls and Ritter (2013) sum up the camps like this:

> The first camp supports traditional teacher preparation in approved degree programs within colleges of education. In their course of study, teaching candidates learn

> educational theory, receive pedagogical training, and have practical classroom experience…[T]he other camp advocates for a more direct path to the classroom, with a focus on content knowledge. According to this strategy, teachers would typically have a degree in the subject they teach, but they have much less classroom experience. This group prefers alternative programs, such as Teach for America, which seeks highly talented individuals with strong content knowledge, provides a shortened training period, usually six weeks, and then places teachers in classrooms. (p. 29)

Generally speaking, the first camp wants to restrict access to the profession by making sure teachers have the proper training before they enter the classroom.

In my estimation, the goal of any teacher preparation program or certification route should be twofold; to prepare teachers for effective service in the classroom; but in order to do that the program must also attract quality applicants. Shuls and Ritter (2013) note, "If we neglect to recognize both aspects of teacher preparation, to prepare and attract, then we do a disservice to the profession and to students" (p. 32). The fact of the matter is that the traditional method of preparing teachers for the classroom fails this test in two regards, quantity and quality.

There are many teachers entering the teaching field through the traditional licensure route. However, these programs are not producing enough teachers in specific subjects or enough teachers willing to work in specific locations (Ingersoll 2011). Rural areas and urban centers routinely have difficulty staffing their classrooms with qualified teachers and there are often shortages in math, science, and special education. I will discuss this problem a bit further in the section on reforming teacher pay, but here it is simply important to note that the traditional licensure route cannot meet the demands of the market in terms of quantity.

The second way the traditional pathway is failing to "prepare and attract" is in terms of the quality of individuals entering the profession. Let me be clear, there are many great and wonderful teachers entering via the traditional path. Nevertheless, it cannot be denied that, on average, individuals entering the profession through traditional pathways tend to perform lower on standardized exams than other college majors. Podgursky, Monroe, and Watson (2004), note "High-ability college graduates are less likely to teach in public schools and, if they do, are more likely to leave after a few years" (p. 515-516). In their data of Missouri college students, they note that only 10.9% of male and 8.8% of female elementary teachers scored a 27 or above on their ACT, compared to 28.9% and 22.9% of non-teachers. The average ACT score for teachers is a full point lower than it is for non-teachers for both genders. In addition to the evidence from ACT scores, it is becoming abundantly clear that alternatively certified teachers score higher on licensure exams. This is true in New York, Florida, and here in Arkansas.

All of this indicates that the traditional pathway to the classroom is not doing an adequate job of attracting high performing teachers to the field. One of the potential obstacles to recruiting teachers through the traditional path is the lack of options they would have with an education degree. When an individual gets a degree as an elementary school teacher they have few options in the marketplace with that degree. The preparation to become a teacher does little to prepare graduates for other careers. This is particularly troubling when fewer people work one job for their entire life. Hess (2009) writes, "Recruiting new college graduates for teaching positions made sense 40 years ago, when the typical graduate could expect to hold just five jobs in an entire career. Today, graduates may have held four jobs by age 30" (p. 36). For all of these reasons, it is important to consider other opportunities to attract and train teachers for the classroom from outside the traditional route.

Lawmakers have expanded routes to the classroom, most likely out of necessity for teachers in shortage areas. Now there are a number of routes that an individual can take to enter the profession that do not require a bachelor's degree in education. However, these programs do not go far enough. Alternative certification routes, as we have conceived of them in Arkansas, are mostly designed for career switchers. That is to say, individuals who wish to become a teacher still face significant barriers to entry. The average person, who may wish to teach part time, fill in a course, or teach for a few years most likely will not choose to pursue a path in teaching because of the obstacles they face. Teach for America, and programs like it, are the exception. These programs attract highly capable individuals and get them to commit to the classroom for two years, with remarkable success. The difference between these programs and our non-traditional licensure program is that TFA creates a pathway for individuals looking for a career, a mission field, or even a trial basis. What we need are pathways designed to attract highly capable people, for whatever reason, not just career switchers.

One way to open the teaching field up to more prospective teachers is to decentralize the certification process from the state to the individual school district. In this type of system, school districts would be free to hire whomever they felt was most qualified for the position. If they wanted to hire a full-time, traditionally certified teacher to teach their high school math courses, they could. But if they wanted to hire a local engineer to teach calculus, and a banker to teach personal finance, and a bright young graduate to teach algebra, they could do that also. A new hire for a school district would receive a provisional teaching certificate. After that teacher had demonstrated their competence in the classroom, the district could sponsor their certification.

Some might argue that removing state requirements for certification would damage the teaching profession. That seems unlikely to me. School administrators want to hire qualified

teaching applicants. When their teachers are highly competent, it makes the administrators look good. Administrators would still be free to hire traditionally trained teachers, but they would also have greater ability to choose applicants who do not have a degree in education. From the simple law of large numbers, having a larger applicant pool to choose from means school leaders would have the ability to hire better teachers, assuming they could identify quality. Moreover, this type of system incentivizes teacher preparation programs to ensure that they are truly equipping their teachers for the classroom; otherwise, their teachers will not be able to find jobs.

We simply cannot plan, train, and organize a high quality teacher workforce as some would like to do. Teacher training may be somewhat beneficial, but when tough rules and regulations are put in place it simply has the effect of limiting the labor pool and barring potentially great teachers from entering the profession. A system such as I have described, where school leaders can choose the candidate they believe is the best, will increase the labor pool and improve teacher quality.

Use Teacher Testing Wisely

It is quite clear that there is a relationship between a teacher's performance on licensure exams, or even other standardized tests, and their performance in the classroom. It is also clear that the tests used for licensure screens are imperfect measures of quality. That means they let in teachers who end up being ineffective teachers and they keep out potentially effective ones. I see no reason to keep using licensure exams, at least not in the capacity that we are currently using them.

Licensure exams do provide useful information for administrators to consider in their hiring process. Unfortunately, most administrators do not look at scores on licensure exams when they are making their important staffing decisions (Shuls, Burks, and Ritter 2012). In an

examination of the application documents at 50 randomly selected Arkansas school districts, Shuls, Burks, and Ritter (2012) discovered that only 13% of the districts requested a prospective teacher's scores on the Praxis exams. Applicants were rarely asked for other traditional measures of academic capability, like college honors (22%) or college GPA and transcripts (52%). Rather, most school districts use certification (98%), teaching experience (98%), or references (96%) to assess the capability of a candidate.

If licensure exams are to be required, they should not be used to determine certification; rather, they should be used to inform decision making. This will be particularly important if we decentralize the certification process. Administrators should use these scores, as well as a whole host of other information to make the most informed decision about who they hire.

Link Teacher Pay More Closely to Quality

Experience and advanced degrees are the primary drivers of teachers' salaries across the country, including here in Arkansas. This is problematic for a number of reasons. For starters, pay should be based, at least in part, on some measure of quality. As we have seen here and in previous literature, there is little evidence that experience matters past the first five years. Moreover, there is no clear evidence to suggest that earning a master's degree improves teacher effectiveness.

Not only does the single salary schedule not fully reward quality, there is suggestive evidence that the structure of the teacher pay system contributes to teacher shortages in hard to staff fields like math and science (Shuls 2012). The argument is that individuals with expertise in math and science can command higher salaries outside of teaching. Therefore, they may be less likely to select into teaching and more likely to leave for a higher paying job. Since experience and master's degrees are not clear measures of quality and there is reason to believe the single

salary schedule contributes to teacher shortages in math and science, it follows that school districts should consider restructuring teacher pay.

Let me be clear, though I believe teacher pay should be reformed, I do not believe these changes should be mandated at the state level. Each school district is uniquely positioned in the state so that one policy to rule them all would not have the effect of improving teacher quality in each district. Nevertheless, there are numerous changes to teacher pay that an individual school district could undertake that would help them attract and retain the teachers they need.

One modest change that school administrators may consider is simply reshaping the existing salary schedule. In a study linking the structure of salary schedules to school level student performance, Grissom and Strunk (2012) note that school districts that frontload their salary structure tend to perform better. A typical salary schedule might start at $35,000 and then give percentage raises for each year. This has the effect of giving larger raises, in terms of real dollars, to more veteran teachers. By simply changing the slope of the salary schedule to provide larger raises in the first few years, school districts may have greater ability to attract bright young individuals. Furthermore, they may be able to retain those teachers within the first five years, a time when many teachers leave the field. In this type of system an individual may make the same amount over the course of their lifetime, but they would make more in the early years when money is typically more of a factor in job retention. This type of change would also do a better job of aligning teacher pay to teacher quality, since teachers improve more in the first few years than in later years.

While changing the structure of the salary schedule may help improve overall teacher quality, I do not believe it addresses the problem of shortages in specific teaching fields. To address this problem, school leaders may want to consider one of the strategies outlined by Shuls

(2012). Districts could place high need teachers on a higher step on an existing salary schedule, they could create separate salary schedules, or they could pay each individual based on their specific needs/school resources at the time. There is no reason that a school district must have one salary schedule or that all teachers must be paid on the same scale. In most fields, individuals with different skill sets are paid varying rates based on their skills and market options. Even in higher education, professors of some subjects make considerably more than professors of other subjects (College and University Professional Association for Human Resources 2012). Since math and science teachers may have more options, it may make sense to place these individuals on a higher paying salary schedule.

Again, paying teachers of shortage areas more should not be mandated from the state. In a conversation with one rural superintendent, he told me that his district had no problem finding math and science teachers, but they could not find foreign language teachers. Thus, the decision for what types of teachers should be paid more should be left to each district, since they know best what types of teachers they have difficulty attracting or retaining.

The plans outlined so far do a better job of linking teacher pay to teacher quality and market competition. However, they do not fully address the issue of linking a teacher's performance in the classroom to their salary. The best way to do this would be to develop a system for performance pay. Ritter and Jensen (2010) outline some important aspects that districts should consider when developing a merit pay system. They suggest that teacher merit pay plans should properly incentivize teachers in productive ways. That means the plan should have a reward that actually motivates teachers to improve; a small reward is unlikely to have much impact. The plan should also encourage teachers to collaborate and work together. Ritter

and Jensen (2012) suggest not having a fixed pot of money, where one teacher's success is another teacher's failure.

Though performance pay seems like an effective strategy for improving teacher quality, the results from actual implementation have not been as promising. As Buck and Greene (2011) note, the "research on this topic within the United States is sparse and results are mixed" (p. 28). They suggest this is because teachers and their unions often block, dilute, or co-opt good performance pay plans. This is an important factor for school districts to consider. Merit pay plans are often difficult to implement. They require additional resources and buy-in from faculty. In some school districts the opposition may simply be too much to overcome and they may wish to opt for one of my prior suggestions.

Whether a school district changes the shape of their salary schedule, implements multiple schedules, uses performance pay, or another method to attract and retain high-performing teachers, true merit pay and improvement in teacher quality will come from removing low-performing teachers. Jay Greene calls this "true merit pay" (2012). He says, "True merit pay—the kind of compensation for job performance found in most industries—provides effective employees with continued employment and regular raises while ineffective workers lose their jobs."

Unfortunately, many school leaders fail to even recognize teacher quality. In a 2009 study of teacher evaluations in 12 large school districts, Weisberg et al. (2012) found that there was little variation in the ratings of teachers. Almost all teachers received high marks on their evaluation, with very few receiving the lowest rating. They concluded that administrators treat teachers like widgets, where one could easily replace the other. Similarly, Jacob, Vidyarthi, & Carroll (2012) found that high-performing teachers were encouraged to keep teaching and

retained by their school district at similar rates as low-performing teachers. If any changes are to be made to improving teacher compensation, we must start by recognizing that teachers vary in quality and their skill sets vary in market value.

Conclusion

Teacher quality is an important issue. In this dissertation, I have explored the relationship between observable teacher characteristics and teaching effectiveness, as measured by value-added student achievement. This study was conducted using data from Arkansas, a state that is much different from other states and cities where this type of analysis has been conducted. There is reason to believe the results from other states may be markedly different from Arkansas because of the quality of individuals attracted to the teaching profession. Yet, the results presented here are in harmony with the existing body of research.

Taken together, the results here and in the broader research indicate that current policies regarding teacher certification and teacher pay are not the most effective way to attract and retain teachers in classrooms. I have outlined here three ways policymakers and school leaders can improve the teacher workforce by creating a system that attracts high-performing individuals and allows schools to retain those individuals by being more flexibly in their pay structures. On this matter Hess (2009) states, the goal should be "to recruit the most promising talent and then foster a more flexible, rewarding, and performance-focused profession" (p. 39). I agree.

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and Student Achievement in the Chicago Public High Schools. *Journal of Labor Economics*, *25*(1).

Angus, D. L. (2001).*Professionalism and the public good: A brief history of teacher Certification.* Washington, DC: Thomas B. Fordham Foundation.

Arkansas Department of Education. (2012). Critical academic licensure shortage areas. Retrieved on February 19, 2013 from http://www.teacharkansas.org/pdfs/2012-2013%20Critical%20Shortage%20Areas.pdf

Kimbrell, T. (2010). Protocol for the review and approval of programs of study leading to educator licensure or endorsement in Arkansas. *Arkansas Department of Education.* Retrieved March 9, 2013 from http://www.adhe.edu/SiteCollection Documents/ AcademicAffairsDivision/Education%20Program%20Protocol.pdf

Ballou, D., Sanders, W. & Wright, P. (2004). Controlling for student background in value-added assessments of teachers. *Journal of Educational and Behavioral Statistics*, 29(1), 37-65.

Ballou, D., & Podgursky, M. (2000). Reforming teacher preparation and licensing: What is the evidence?. *The Teachers College Record*, *102*(1), 5-27.

Betts, J., Zau, A. C., & Rice, L. A. (2003). *Determinants of student achievement: New evidence from San Diego.* San Francisco: Public Policy Institute of California. Retrieved March 9, 2013 from http://www.ppic.org/content/pubs/report/r_803jbr.pdf

Boyd, D.J., Grossman, P. L., Lankford, H., Loeb, S., 7 Wyckoff, J. (2006). How changes in entry requirements alter the teacher workforce and affect student achievement. *Education Finance and Policy*, 1(2), 176-216.

Boyd, D., Lankford, H., Loeb, S., Rockoff, J., & Wychoff, J. (2008). The narrowing gap in New York City teacher qualifications and its implications for student achievement in high poverty schools. *Journal of Policy Analysis and Management*,27(4), 793-818.

Buck, S. & Greene, J. P. (2011). Blocked, diluted, and co-opted. *Education Next*, 11(2), 26-31.

Cantrell, S., Fullerton, J., Kane, T. J., & Staiger, D. O. (2008). National board certification and teacher effectiveness: Evidence from a random assignment experiment. *National Bureau of Economic Research*, Working Paper 14608. Retrieved March 9, 2013 from http://www.nber.org/papers/w14608

Cavalluzzo, L. C. (2004). Is National Board Certification An Effective Signal of Teacher Quality? Alexandria, Virginia: The CNA Corporation. Retrieved March 9, 2013 from http://www.nbpts.org/userfiles/file/final_study_11204_d_-_cavalluzzo_-_cna_corp..pdf

Clark Tuttle, C., Anderson, T., & Glazerman, S. (2009). ABCTE teachers in Florida and their effect on student performance. *Mathematica Policy Research Inc.* 6215-400. Retrieved March 9, 2013 from http://www.policyarchive.org/handle/10207/bitstreams/20544.pdf

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *The Journal of Human Resources.* 41(4). pp.778-820.

Clotfelter, C. T., Ladd, H. F., Vigdor, J. L. (2007). How and why do teacher credentials matter for student achievement? *National Center for Analysis of Longitudinal Data in Education Research*, working paper 2. Retrieved March 9, 2013 from http://www.caldercenter.org/PDF/1001058_Teacher_Credentials.pdf

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *The Journal of Human Resources,* 45(3), 665-681.

Cochran-Smith, M. (2005). The new teacher education: For better or for worse?. *Educational Researcher, 34*(7), 3-17.

Coleman, J.S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D. & York, R. L. (1966). *Equality of Educational Opportunity*. Washington, DC: US Department of Health, Education & Welfare. Office of Education (OE-38001 and supp.) Retrieved March 9, 2013 from http://digital.tcl.sc.edu/cdm/search/collection/eeo

College and University Professionals Association for Human Resources (2012). *National Faculty Survey*. Retrieved from March 9, 2013 from http://oipr.depaul.edu/Market_Analytics/documents/NFSS12ExecutiveSummary.pdf

Constantine, J., Player, D., Silva, T., Hallgren, K., Grider, M., Deke, J., &Warner, E. (2009). An evaluation of teachers trained through different routes to certification. (NCEE 2009- 4043). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved March 9, 2013 from http://ies.ed.gov/ncee/pubs/20094043/index.asp

Croninger, R. G., Rice, J. K., Rathbun, A., & Nishio, M. (2003). Teacher qualifications and first grade achievement: A multilevel analysis. *Center for Education Policy and Leadership*. Document OP-03-2. Retrieved March 9, 2013 from http://www.education.umd.edu/EDPS/CEPAL/pdf/oc-papers/CEPAL-OP-03-2.pdf

Daly, W. J. (1998). Peter's Medicine: Lessons from the 13th Century. *Transformations of the American Clinical and Climatological Association*. 109. 243-258. Retrieved March 9, 2013 from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2194354/

Darling-Hammond, L. (1997). *Doing What Matters Most: Investing in Quality Teaching*.

National Commission on Teaching & America's Future. Retrieved March 9, 2013 from http://www.eric.ed.gov/PDFS/ED415183.pdf

Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the Evidence. *Educational Evaluation and Policy Analysis*. 23(1), 57-77.

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Vasquez Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives*. 13(42), 1-48. Retrieved March 9, 2013 from http://epaa.asu.edu/ojs/oldepaa/redirect/?v=13&n=42

Dee, T. (2004). Teachers, race, and student achievement, in a randomized experiment. *The Review of Economics and Statistics*, 86(1), 195-210.

Educational Testing Service. (2013). Praxis I overview. Retrieved on February 19, 2013 from http://www.ets.org/praxis/about/praxisi

Ehrenberg, R. G. & Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from High School and Beyond. *Economics of Education Review*, 13(1), 1-17.

Feistritzer, C. E. (2011). Profile of teachers in the U.S. 2011. *National Center for Education Information.* Retrieved from http://www.edweek.org/media/pot2011final-blog.pdf

Feistritzer, C. E. and Chester, D. T. (2000). *Alternative teacher certification. A state-by-state analysis*. Washington, D. C. National Center for Education Information.

Feng, L. & Sass, T. (2010). What makes special-education teachers special? Teacher training and achievement of students with disabilities. *National Center for Analysis of Longitudinal Data in Education Research*. Working paper 49. Retrieved March 9, 2013 from http://www.caldercenter.org/upload/CALDERWorkPaper_49.pdf

Ferguson, R. F. & Ladd, H. F. (1996) How and why money matters: An analysis of Alabama schools. In Ladd, H.F. (Ed.), *Holding schools accountable: Performance-based reform in education*. Washington, DC.: The Brookings Institution.

Friedman, M. & Kuznets, S. (1945). Income from Independent Professional Practice. *National Bureau of Economic Research*, New York. Retrieved March 9, 2013 from http://papers.nber.org/books/frie54-1

Gellhorn, W. (1976). The abuse of occupational licensing. *The University of Chicago Law Review*, 44(1), 6-27.

Glazerman, S., Mayer, D., & Decker, P. (2006). Alternative routes to teaching: The impacts of Teach for America on student achievement and other outcomes. *Journal of Policy Analysis and Management*, *25*(1), 75-96.

Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources*, 42(4), 765-794.

Goldhaber, D. (2011). Licensure: Exploring the value of this gateway to the teacher workforce. In Hanushek, E., Machin, S. J., & Woessman, L. (Eds), *Handbook of the Economic of Education, Volume 3*. San Diego: North-Holland.

Goldhaber, D. & Anthony, E. (2007). Can teacher quality be effectively assessed? National board certification as a signal of effective teaching. *The Review of Economics and Statistics*, 89(1), 134-150.

Goldhaber, D. D. & Brewer, D. J. (1997a). Evaluating the effect of teacher degree level on educational performance. In Developments in school finance 1996, edited by William Fowler, 197-210. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Goldhaber, D. D. & Brewer, D. J. (1997b). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *Journal of Human Resources*, 32(3), 505-23.

Goldhaber, D. & Brewer, D. (2000) Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*. 22(2), 129-145.

Greene, J. P. (2012). In Chicago—phony merit pay is dead, long live true merit pay. *Jay P. Greene's Blog*, Retrieved March 9, 2013 from http://jaypgreene.com/tag/merit-pay/

Grissom, J. & Strunk, K. (2012). How should school districts shape teacher salary schedules? Linking school performance to pay structure in traditional compensation schemes. *Educational Policy*, 26(5), 663-695.

Hanushek, E. (2010). The economic value of higher teacher quality. *National Bureau of Economic Research*, Working paper 16606, Cambridge, MA: National Bureau of Economic Research. Retrieved March 9, 2013 from http://www.nber.org/papers/w16606

Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). The market for teacher quality. *National Bureau of Economic Research*, Working paper 11154. Retrieved March 9, 2013 from http://www.nber.org/papers/w11154

Hanushek, E. A., Kain, J. F., Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of Human Resources*, 39, 326-354.

Harper, C. A. (1939). *A century of public teacher education: The story of the state teachers colleges as they evolved from the normal schools*. Washington D.C.: Hugh Birch-Horace Mann Fund for the American Association of Teachers Colleges. Retrieved from http://archive.org/details/centuryofpublict00harp

Harris, D. N. & Sass, T. R. (2011). Teacher training, teacher quality, and student achievement. *Journal of Public Economics*, 95(2011), 798-812.

Hess, F. M. (2009). How to get the teachers we want. *Education Next*, 9(3), 34-39.

Ingersoll, R. (2011). Do we produce enough mathematics and science teachers? *Phi Delta Kappan*, 92(6), 37-41.

Jacob, A., Vidyarthi, E., & Carroll, K. (2012). The Irreplaceables: Understanding the Real Retention Crisis in America's Urban Schools. *TNTP*. Retrieved March 9, 2013 from http://tntp.org/irreplaceables

Kane, T. J., Rockoff, J. E., & Staiger, D. O. (2008). What does certification tell us about teacher effectiveness? Evidence from New York City. *Economics of Education Review*, 27(6), 615-631.

Kleiner, M. M. (2000). Occupational licensing. *The Journal of Economic Perspectives*, 14(4), 189-202.

Kleiner, M. M. & Krueger, A. B. (2010). The prevalence and effects of occupational licensing. *British Journal of Industrial Relations*. 48(4), 676-687.

Krieg, J. M. (2006). Teacher quality and attrition. *Economics of Education Review*, 25(1), 13-27.

National Center for Education Statistics. (2008). *Schools and Staffing Survey*. Retrieved March 9, 2013 from http://nces.ed.gov

National Center for Education Statistics. (2010a). Enrollment in public schools. *Common Core of Data*. Retrieved March 9, 2013 from http://nces.ed.gov

National Center for Education Statistics. (2010b). Enrollment in private schools. *Private School Universe Survey.* Retrieved March 9, 2013 from http://nces.ed.gov

Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237-257.

Obama, B. (2009). Remarks by the president in Arnold, Missouri Town Hall. *The White House Office of the Press Secretary*. retrieved May 9, 2011 from http://www.whitehouse.gov/the-press-office/remarks-president-arnold-missouri-town-hall

O'Donnell, P. (2010). Old deluder Satan law. In Hunt, T., Carper, J., Lasley, T., & Raisch, C. (Eds.), *Encyclopedia of Educational Reform and Dissent*. (pp. 676-677). Thousand Oaks, CA: SAGE Publications, Inc.

Podgursky, M. (2005). Response to "Does teacher preparation matter?" *National Council on*

Teacher Quality. Retrieved March 9, 2013 from http://www.nctq.org/nctq/research/1114011044390.pdf

Podgursky, M. Monroe, R., & Watson, D. (2004). The academic quality of public school teachers: An analysis of entry and exit behavior. *Economics of Education Review,* 23(2004), 507-518.

Ravitch, D. (2003). *A brief history of teacher professionalism*. White House conference on preparing tomorrow's teachers. Retrieved March 9, 2013 from http://www.udel.edu/educ/whitson/897s05/files/ravitch_Teacher_Professionalism.htm

Ravitch, D. (2010). In Need of a Renaissance. *American Educator, 34*(2), 11.

Reese, W. J. (2011). *America's public schools: From the common school to "No Child Left Behind"*. Baltimore, MD: The Johns Hopkins University Press.

Ritter, G. W. & Shuls, J. V. (2012). If a tree falls… *Phi Delta Kappan*, 94(3), 34-38.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica,* 73(2) 417-458.

Rockoff, J. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *The American Economic Review,* 94(2), 247-252.

Rockoff, J. E., Jacob, B. A., Kane, T. J., & Staiger, D. O. (2008). Can you recognize an effective teacher when you recruit one? *National Bureau of Economic Research working paper series*, Working paper 14485. Retrieved March 9, 2013 from http://www.nber.org/papers/w14485

Rottenberg, S. (1962). The economics of occupational licensing. In Lewis, H. Gregg, ed. 1962. *Aspects of Labor Economics*. Princeton: Princeton University Press. Retrieved May 9, 2011 from www.nber.org/books/univ62-2

Rottenberg, Simon. 1980. "Introduction," in *Occupational Licensure and Regulation*. S. Rottenberg, ed. Washington, American Enterprise Institute, pp. 1-10.

Sanders, W. L., Ashton, J. J., & Wright, S. P. (2005). Comparison of the Effects of NBPTS Certified Teachers with Other Teachers on the Rate of Student Academic Progress. Technical Report. SAS Institute. Retrieved March 9, 2013 from http://www.nbpts.org/resources/research/browse_studies?ID=15

Sass, T. R. (2006). Charter schools and student achievement in Florida. *Education Finance and Policy,* 1(1), 91-122.

Sass, T. R. (2011). Certification requirements and teacher quality: A comparison of alternative routes to teaching, *Working paper.* Retrieved March 9, 2013 from http://www2.gsu.edu/~tsass/Alternative%20Certification%20and%20Teacher%20Quality%2011.pdf

Shuls, J. (2012). The salary straitjacket: The pitfalls of paying all teachers the same. *Show-Me Institute Essay*, Retrieved March 9, 2013 from http://www.showmeinstitute.org/publications/essay/education/847-single-salary-schedules.html

Shuls, J. V., Burks, S. M. & Ritter, G. W. (2012). What do schools look for in prospective teachers? An analysis of teacher application documents. *Office for Education Policy, Arkansas Education Report*, 9(3). Retrieved March 9, 2013 from http://www.officeforeducationpolicy.org/wp-content/uploads/2013/01/9_3_What_do_Schools_Want.pdf

Shuls, J. V. & Ritter, G. W. (2013). Teacher preparation: Not an either/or. *Phi Delta Kappan*, 94(7), 28-32..

Smith, A. (1863). *An inquiry into the nature and causes of the wealth of nations*. A. and C. Black.

Todd, P.E., & Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485), F3–F33.

Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics*, 33(1), 1-67.

United States Census Bureau. (2010). Urban and rural universe: Housing units.

United States Department of Education. (2008). The final report of the national mathematics advisory panel. Retrieved March 9, 2013 from http://www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

U.S. News and World Report. (2013). National University Rankings. Retrieved on February 19, 2013 from http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/national-universities

Walsh, K. (2001). *Teacher certification reconsidered: Stumbling for quality.* The Abell Foundation. Retrieved March 9, 2013 from http://www.abell.org/pubsitems/ed_cert_1101.pdf

Walsh, K. & Jacobs, S. (2007). Alternative certification isn't alternative. *National Council on Teacher Quality.* Retrieved March 9, 2013 from http://news.heartland.org/sites/all/modules/custom/heartland_migration/files/pdfs/22264.pdf

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect. *Brooklyn, NY: The New Teacher Project*. Retrieved March 9, 2013 from http://widgeteffect.org/

West, M. R., Chingos, M. M. (2009). Teacher effectiveness, mobility, and attrition in Florida. In M. G. Springer (Ed.), *Performance incentives: Their growing impact on American K-12 education* (pp. 251–271): Washington, DC: Brookings Institution Press.

Winer, B. J. (1971). *Statistical principles in experimental design*. Second Edition. McGraw-Hill, New York.

Xu, Z., Hannaway, J., & Taylor, C. (2011). Making a difference? The effects of Teach for America in high school. *Journal of Policy Analysis and Management*. 30(3), 447-469.