SELECTING AND PREPARING TEACHERS AND SCHOOL LEADERS TO IMPROVE EDUCATIONAL OUTCOMES

Kevin C. Bastian

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Public Policy.

Chapel Hill
2013

Approved by:

Gary Henry, Ph.D.

Charles Thompson, Ph.D.

Douglas Lauen, Ph.D.

Jeremy Moulton, Ph.D.

Bill McDiarmid, Ph.D.

Charles Clotfelter, Ph.D.

ABSTRACT

KEVIN C. BASTIAN: Selecting and Preparing Teachers and School Leaders to Improve Educational Outcomes
(Under the direction of Gary Henry)

In the following three studies we explore ways to improve the quality of school-based personnel by identifying characteristics and training of effective teachers and principals. First, to enhance the selection and hiring of teachers into preparation programs and/or school districts, we examine whether teachers' non-cognitive characteristics predict teacher value-added and evaluation ratings. This work builds upon research in psychology and economics which shows the impact of individuals' non-cognitive attributes on outcomes of interest. Here, evidence indicates that teachers' non-cognitive characteristics significantly influence student achievement gains and predict higher evaluation ratings. Due to the select sample of teachers examined in this analysis—Teach For America corps members—this work represents a "proof of concept" and calls for continued research on a more representative sample of teachers.

Second, to better inform the choices made by states and districts in staffing schools, we question whether effectiveness differences exist between teachers traditionally prepared in-state versus out-of-state, and if so, we test three research-based hypotheses to explain differences. Overall, the findings indicate that out-of-state prepared teachers are significantly less effective than in-state prepared and alternative entry teachers in North Carolina elementary schools. Further, out-of-state prepared teachers' lack of familiarity with the state's educational environment and attrition

patterns—high rates of turnover coupled with the attrition of less effective teachers—help explain their performance. This suggests policy mechanisms to increase the in-state prepared teaching population and improve the quality of out-of-state prepared teachers.

Finally, to contribute to the nascent principal quality research agenda, we detail the characteristics of first-time principals and the schools that hire them. Further, we question whether individual principal characteristics or those of the environments in which they previously worked are associated with student achievement gains. Descriptively, we find that a majority of first-time principals are "homegrown"—promoted from within the district—and evidence that first-time principals sort into schools based on observed characteristics. Our strongest value-added results indicate that early-career principals who served in high value-added schools as assistant principals promote greater student achievement gains. This suggests that principals learn key aspects of effective school leadership during their assistant principal experience.

ACKNOWLEDGEMENTS

I came to North Carolina in the summer of 2006, fresh out of college and ready to make a difference as a Teach For America corps member. As an undergraduate I had studied the Civil Rights Movement and naturally, saw educational equity as the civil rights issue of today. During my two years teaching sixth grade language arts I learned a lot about myself and what I wanted. Most importantly, I learned that teaching was not my long-term career—I was only an average teacher at best—but that education, especially finding ways to systematically improve educational quality and outcomes for disadvantaged students, was a real passion. This led me to public policy and the belief that research and evidence-based policymaking plays a vital role in systematic change. Along this path, as I have progressed from a novice graduate student to an independent education researcher, I have learned and received support from many people who I wish to thank for helping me get to where I am today.

I am deeply grateful to my advisor, Gary Henry, who has been a terrific mentor and research partner. I am particularly indebted to Gary for the research leadership opportunities he provided me—directing research projects, framing and writing articles, presenting results to stakeholders—that have significantly shaped my development. He was quick to acknowledge high-quality work and listened to my ideas and opinions. With him I always felt like a colleague. Gary helped guide my research focus to teacher

and principal quality and more than anyone else, taught me how to think like a researcher. He has been the driving force to shape me into the researcher I am today.

I consider myself especially fortunate to have had a second mentor, Charles Thompson, throughout my graduate school tenure. Given our mutual affection for southern history and high-quality BLT sandwiches, it was always easy to hit it off with Charles. When it came to research Charles taught me the value of asking questions, challenged me to express ideas with greater clarity, and modeled how to make research findings graspable for the audience. I am deeply thankful to Charles for the faith and confidence he has always shown in me and for his commitment to improve education in North Carolina.

All of the former and current members of the Education Policy Initiative at Carolina (EPIC) helped created a fun, supportive, and educative research experience for me. I started at EPIC with limited data management and quantitative analysis skills; through hands-on mentorship and trust in me to successfully complete important research projects I developed as a researcher. I want to especially thank: Kelly Purtell, who was never too busy to answer a question, talk Carolina basketball, or give me a hard time; Dave Kershaw, who helped me understand our datasets and was always available to talk through a puzzling result; Elizabeth D'Amico, who often single-handedly kept EPIC running smoothly and was always ready to listen; Julie Marks, with whom I shared research leadership experiences that were always enjoyable and educative; and Kevin Fortner, who in my early years at EPIC received many phone calls asking for assistance and always managed to both solve the problem and make me more independent at the

same time. I feel fortunate to have learned from all of you, I look forward to continuing research projects with you, and I am happy to call you all friends.

I would also like to acknowledge my broader research community in North Carolina. I thank all of my committee members, including Doug Lauen, Jeremy Moulton, Bill McDiarmid, and Charles Clotfelter for providing guidance and support throughout my dissertation. I am honored to have worked with each of you and I look forward to future research collaboration. I also wish to thank Alisa Chapman with the University of North Carolina General Administration and the University of North Carolina Council of Education Deans for being truly great research partners. Our research has improved the teacher preparation programs at UNC system schools and I believe the best is yet to come.

I must also acknowledge my amazing graduate student cohort of Rod Rose, Shanyce Campbell, Jade Marcus, and Shelley Golden. Together, we made it through math camp, classes, and all our departmental exam requirements. Along the way we enjoyed lunches, ran up the steepest hills in Chapel Hill, worked on research projects, celebrated accomplishments and holidays, and shared in landmark moments in each other's lives. I believe graduate school is an experience best done with a cohort and I could not have asked for a better one. I look forward to seeing all the incredible things you do.

Beyond research, I am exceedingly grateful to friends that made North Carolina such a wonderful home. I am happy to thank Michael and Elisha Johnson, Rod Rose and Lisa McCaskill, and Craig and Nisha Cavanaugh. Whether it was time on the golf course, running on a trail, or sharing a good meal you all provided many special

memories. Most of all I thank Kiani and Matthew Gardner—I cannot think of North Carolina without thinking of our trivia nights, time on the lake, and nearly constant laughter.

As I conclude my graduate school career, I am deeply grateful for the encouragement and love I have received from my family. I would like to thank my in-laws, Wayne and Chris Itano, for their support and the incredible examples they have set, both in their academic careers and in the family they created. Words cannot express my gratitude for my parents, Bob and Debbie Bastian. They have given me unconditional love and taught me about hard-work, caring, and pursuing your passions. I see so much of who I am in them. Finally, I thank my wife and best friend, Michelle Itano. Every day you challenge me to be better, make me laugh, and provide love and support. You are my partner on this journey and I look forward to all that lies ahead for us.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

OVERVIEW

The provision of high-quality public education entails clear, substantial benefits for both individuals and society. Individual wage returns to another year of education are approximately ten percent, larger than the relative returns from other investments (Card, 1999; Harmon, Oosterbeek, & Walker, 2003; Psacharopoulos & Patrinos, 2004). Beyond wages, educational attainment is also associated with a number of other desired, individual and societal outcomes: improved health status and decreased mortality rates; reduced rates of unemployment and incarceration; improved child outcomes, including future educational attainment; increased tax contributions; and greater levels of civic participation (Adams, 2002; Dee, 2004; Jencks, 1972; Lleras-Muney, 2005; Lochner & Moretti, 2004; Ross & Wu, 1995; Rouse, 2005). Quite simply, in an increasingly competitive global economy, sizable economic incentives exist for policymakers to create education systems that promote greater achievement.

These benefits of education make the performance of the public education system in the United States even more troubling. Results from the most recent Program for International Student Assessment (PISA) show that, relative to students in other Organization for Economic Cooperation and Development (OECD) countries, American students rank fifteenth, thirty-first, and twenty-third, respectively, in reading, mathematics, and science achievement (Walker, 2011). While the quality of public education in the United States is one factor that accounts for these rankings, recent research also suggests that (1) the high concentrations—relative to other OECD

countries—of economically disadvantaged students in the United States and (2) the large achievement gaps between more and less-affluent students in the United States explains a substantial portion of these international performance differences (Carnoy & Rothstein, 2013). To quantify these achievement gaps, results from the 2011 4th grade mathematics National Assessment of Educational Progress (NAEP) show that, on average, students eligible for free school lunches scored 23 points below—0.79 standard deviation units—their non-eligible peers (National Center for Education Statistics, 2012). Further exacerbating these performance disparities in the United States is the inequitable distribution of educational resources—peers, teachers, and funding—to academically at-risk students (Bastian, Henry, & Thompson, 2012; Burke & Sass, 2008; Clotfelter, Ladd, & Vigdor, 2005; Downes & Stiefel, 2008; Lankford, Loeb, & Wyckoff, 2002). Overall, the effects of these achievement and inequality statistics threaten societal and economic development and present a significant impetus for policy action.

To determine how to best respond to these education challenges, policymakers can rely on quantitative and qualitative research evidence to identify promising policy directions. Here, research findings clearly indicate that non-school factors strongly influence education outcomes. Dating from the time of the Coleman Report, research has shown that family and demographic factors, such as parental education and involvement and socio-economic status, explain a substantial portion of the variance in student achievement (Coleman et al., 1966; Jencks, 1972; Sewell & Hauser, 1972). Therefore, although not directly linked to school quality, policies targeted at improving the educational and socio-economic status of disadvantaged families and the environments in

which they live may have long-term, beneficial effects on academic outcomes (Ladd, 2012).

Outside these background characteristics, research findings also clearly demonstrate the effects of early-childhood education and school-related variables. Participation in early-childhood education programs, such as Perry Preschool, Head Start, or universal pre-kindergarten result in higher levels of educational attainment, salaries, and homeownership and lower rates of welfare assistance and criminal behavior (Heckman, 2006). Once in school, organizational factors, such as the assignment of students to academic tracks, performance-based accountability systems, and school governance structures significantly influence educational outcomes (Bidwell & Kasarda, 1980; Bifulco & Ladd, 2006; Booher-Jennings, 2005; Carnoy & Loeb, 2002; Ladd & Lauen, 2010). More important than these organizational factors, however, is the quality of classroom teachers. Dating again from the Coleman Report, research definitively evinces that (1) teacher effectiveness is the school variable explaining the most variation in student achievement and (2) there is substantial variation in teacher effectiveness (Aaronson, Barrow, & Sander, 2007; Nye, Konstantopoulos, & Hedges, 2004; Rockoff, 2004). Estimated teacher effects range from 0.11 to 0.36 standard deviation units in mathematics and from 0.08 to 0.26 standard deviation units in reading; quantified differently, the effect of a highly effective teacher is equivalent to 7.5 and 3 months of additional student learning in mathematics and reading, respectively (Gates Foundation, 2010; Hanushek & Rivkin, 2010). Recent research also suggests that highly effective teachers have economically substantial and long-lasting effects. High value-added teachers annually generate marginal gains of over $400,000 (present value) in future

earnings for a class of twenty-students, and students assigned to these high value-added teachers are more likely to attend college and higher-ranked colleges, live in higher SES neighborhoods, and save more for retirement (Chetty, Friedman, & Rockoff, 2011; Hanushek, 2011). For policymakers interested in education outcomes, improving teacher quality is a clear direction for action.

While highly effective teachers are vital to student achievement growth, they teach, collaborate, and develop within a school context that is shaped by a school principal (Kennedy, 2010). A principal's leadership helps determine whether students, teachers, and the school as a whole succeed (Branch, Hanushek, & Rivkin, 2012). Within a school principals promote improved academic outcomes through multiple mechanisms: (1) recruiting and retaining high-quality teachers, while facilitating the exit of less effective instructors; (2) articulating a shared school vision, culture, and learning goals; (3) providing instructional leadership and structures to support teaching and learning; and (4) allocating school resources towards desired ends (Eberts & Stone, 1988; Grissom & Loeb, 2011; Jacob & Lefgren, 2008; Robinson, Lloyd, & Rowe, 2008; Thompson, Brown, Townsend, Henry, & Fortner, 2011). Large-scale quantitative studies find that (1) principals matter more in academically disadvantaged environments—there is more variation in principal effectiveness in high-poverty schools—and (2) improving principal effectiveness by one standard deviation is associated with student achievement gains of 0.10 to 0.20 standard deviations and increases in graduation rates of 2.6 percentage points (Branch, Hanushek, & Rivkin, 2012; Coelli & Green, 2011; Dhuey & Smith, 2012).

Overall, research evidence clearly demonstrates that the quality of teachers and school leaders significantly affects student achievement, and importantly, can mitigate the influence of family and demographic factors. To improve educational outcomes policymakers must increase the quality of school-based human capital available to students. This means enacting policies that better recruit/select, prepare, develop, evaluate, distribute, and retain high quality teachers and principals. Within each of these policy areas key questions remain, and therefore, the challenge for education policy researchers is to assemble evidence identifying promising policy practices.

In response to this challenge, I present a three chapter dissertation that contributes to the teacher and principal quality research agenda by: (1) identifying characteristics and training of effective teachers and principals (2) detailing directions for continued, follow-up research and (3) suggesting promising policy practices to promote teacher and principal effectiveness.

In my first dissertation chapter, I focus on teacher recruitment and selection and ask whether there are traits of pre-service teachers that predict future value-added effectiveness and evaluation ratings of teacher competencies. Importantly, this effort moves beyond prior teacher selection research—which generally focused on academic credentials—and examines the influence of teachers' non-cognitive skills and traits, such as leadership, perseverance, and organizational ability. Research in other academic disciplines, principally economics and psychology, shows that non-cognitive characteristics significantly impact outcomes of interest; nascent work in education suggests that non-cognitive traits influence teacher practices and quality (Clingman & Fowler, 1976; Duckworth, Quinn, & Seligman, 2009; Heckman & Rubinstein, 2001;

Mischel, Shoda, & Rodriguez, 1989; Rockoff, Jacob, Kane, & Staiger, 2011). Using data provided by Teach For America (TFA), I find that: (1) in elementary grades teachers' organizational ability strongly predicted students' achievement gains, while in high school, teachers' respect (academic and behavior expectations) for students significantly predicted value-added effectiveness and (2) measures of teacher leadership and motivational ability significantly predicted higher evaluation ratings across multiple teacher competencies. While these results illustrate that non-cognitive characteristics can significantly affect outcomes of interest, the highly select research sample—TFA corps members—necessitates further work to determine whether and how the effects of non-cognitive characteristics generalize to a wider population of teachers. Therefore, this research represents a "proof of concept" and impetus for partnerships between researchers and practitioners (teacher preparation programs and school districts) to gather similar non-cognitive data and test effects. The policy implications of such work are clear, as findings could better equip teacher preparation programs and school districts to: (1) recruit and select/hire prospective teachers and (2) cultivate these non-cognitive characteristics.

My second dissertation chapter transitions from teacher selection to teacher preparation and examines whether there are significant differences in elementary grades teacher effectiveness between those teachers traditionally prepared in-state versus out-of-state. From a research perspective, this work makes an important contribution by questioning whether effectiveness differences exist within the traditionally prepared teacher population. From a policy perspective, this work is particularly relevant in North Carolina, where (1) the state's public schools have seen a 15 percent increase in student

enrollment over the previous decade and (2) approximately 30 percent of the teacher workforce, with higher concentrations in elementary schools, earned teacher preparation degrees in other states. Overall, I find that out-of-state prepared teachers are significantly less effective in elementary grades mathematics and reading than those traditionally trained in-state; out-of-state teachers also significantly underperform alternative entry instructors. In order to determine why out-of-state prepared teachers are less effective I tested three research-based hypotheses: (1) teachers with lower levels of human capital may need to be more mobile (moving across state lines) to find employment (Boyd, Lankford, Loeb, & Wyckoff, 2005; Reininger, 2012); (2) differences in state curricula, standards, and culture may lead to out-of-state prepared teachers having less familiarity with the educational environment of the importing state (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009); and (3) out-of-state prepared teachers who acquire human capital through on-the-job experience may become more competitive for teaching positions back in their state of origin, causing both high rates of teacher attrition and differential teacher attrition (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006). Results from analyses testing these explanations indicate that out-of-state prepared teachers underperform due to their (1) lack of familiarity with the schools, students, and curricula of North Carolina and (2) high rates of teacher turnover, coupled with the differential attrition of less effective out-of-state prepared instructors. These findings suggest that North Carolina could improve the staffing of its public schools and subsequently, student achievement, by (1) increasing the total number of in-state prepared teachers in the workforce (2) providing induction programs to ease out-of-state

prepared teachers' transition into the state and (3) identifying screening and hiring practices to better select out-of-state prepared teachers.

Finally, in recognition of principals' significant effects on student achievement and school outcomes, my third dissertation chapter asks: (1) what are the characteristics of first-time principals? and (2) whether individual characteristics of principals or the environments in which they previously worked are significantly associated with student achievement gains? Given both the high concentration of early-career principals—over one-half of principals have less than 5 years experience—and the adverse academic effects of schools transitioning to a first-time principal, I focus my descriptive and empirical analyses on four cohorts of first-time principals (2006-07 through 2009-10). Descriptively, I find that first-time principals are "homegrown"—a large majority assumes the principalship within the same district in which they once worked as teachers and/or assistant principals—and evidence that better-credentialed first-time principals sort into schools with differing levels of prior academic performance, student body composition, and teacher workforce credentials. Concerning student achievement gains, my estimates are not causal, but rather, identify associations of interest that serve as the foundation for hypothesis generation and future research. Several individual principal characteristics are significantly associated with student achievement, however, my strongest findings concern characteristics of principals' prior work environments, especially (1) a congruence between the level (elementary/middle/high) of the assistant principalship and principalship schools and (2) the effectiveness (value-added) of the assistant principalship school. Each of these findings indicates that beginning principals may learn key aspects of effective school leadership during their assistant principalship

tenure. From a policy perspective, this suggests that school districts' patterns of assigning assistant principals and principals to schools should be carefully considered.

Overall, research findings from the past decade evince a broad consensus: the quality of school-based personnel, especially teachers and principals, significantly influences student achievement. Now, to inform policy and improve outcomes of interest, a major component of education research is focused on answering the following question—what makes teachers and principals effective? It is to this endeavor that this dissertation seeks to make a unique and significant contribution.

REFERENCES

Adams, S. (2002). Educational attainment and health: Evidence from a sample of older adults. *Education Economics, 10*(1), 97-109.

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95-135.

Bastian, K.C., Henry, G.T., & Thompson, C.L. (2013). Incorporating access to more effective teachers into assessments of educational resource equity. In press, *Education Finance and Policy.*

Bidwell, C., & Kasarda, J. (1980). Conceptualizing and measuring the effects of school and schooling. *American Journal of Education, 88*(4), 401-430.

Bifulco, R., & Ladd, H. (2006). The impacts of charter schools on student achievement: Evidence from North Carolina. *Education Finance and Policy, 1*(1), 50-90.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal, 42*(2), 231-268.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management, 24*(1), 113-132.

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

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*(4), 416-440.

Branch, G., Hanushek, E., & Rivkin, S. (2012). Estimating the effect of leaders on public sector productivity: The case of school principals. *NBER working Paper 17803.*

Burke, M., & Sass, T. (2008). Classroom peer effects and student achievement. *Federal Reserve Bank of Boston Working Paper 08-5.*

Card, D. (1999). The causal effect of education on earnings. *Handbook of Labor Economics, 3*(1), 1801-1863.

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis, 24*(4), 305-331.

Carnoy, M., & Rothstein, R. (2013). What do international tests really show about U.S. student performance. *Economic Policy Institute.* Available from: http://www.epi.org/files/2013/EPI-What-do-international-tests-really-show-about-US-student-performance.pdf.

Chetty, R., Friedman, J., & Rockoff, J. (2011). The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood. *NBER Working Paper 17699.*

Clingman, J., & Fowler, R. (1976). The effects of primary reward on the I.Q. performance of grade-school children as a function of initial I.Q. level. *Journal of Applied Behavior Analysis, 9*(1), 19-23.

Clotfelter, C., Ladd, H., & Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review, 24*(4), 377-392.

Coelli, M., & Green, D. (2012). Leadership effects: School principals and student outcomes. *Economics of Education Review, 31*(1), 92-109.

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., et al. (1966). *Equality of Educational Opportunity.* Washington, DC: U.S. Government Printing Office.

Dee, T. (2004). Are there civic returns to education? *Journal of Public Economics, 88*(9-10), 1697-1720.

Dhuey, E., & Smith, J. (2012). How important are school principals in the production of student achievement? Available from: www.sole-jole.org/11034.pdf.

Downes, T., & Stiefel, L. (2008). Measuring equity and adequacy in school finance. In H. Ladd & E. Fiske (Eds.), *Handbook of Research in Education Finance and Policy.* New York: Routledge.

Duckworth, A. Quinn, P., & Seligman, M. (2009). Positive predictors of teacher effectiveness. *The Journal of Positive Psychology, 4*(6), 540-547.

Eberts, R., & Stone, J. (1988). Student achievement in public schools: Do principals make a difference? *Economics of Education Review, 7*(3), 291-299.

Gates Foundation. (2010). Learning about teaching: Initial findings from the Measures of Effective Teaching project. *MET Project Policy Brief.*

Grissom, J., & Loeb, S. (2011). Triangulating principal effectiveness: How perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills. *American Educational Research Journal, 48*(5), 1091-1123.

Hanushek, E., & Rivkin, S. (2010). Using value-added measures of teacher quality. *Calder Institute Policy Brief 9.*

Hanushek, E. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466-479.

Harmon, C., Oosterbeek, H., & Walker, I. (2003). The returns to education: Microeconomics. *Journal of Economic Surveys, 17*(2), 115-156.

Heckman, J.J., & Rubinstein, Y. (2001). The importance of non-cognitive skills: Lessons from the GED testing program. *American Economic Review, 91*(2), 145-149.

Heckman, J.J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900-1902.

Jacob, B., & Lefgren, L. (2008). Can principals identify effective teachers? Evidence on subjective performance evaluation in education. *Journal of Labor Economics, 26*(1), 101-136.

Jencks, C. (1972). *Inequality: A re-assessment of the effect of family and schooling in America.* New York: Basic Books.

Kennedy, M. (2010). Attribution error and the quest for teacher quality. *Educational Researcher, 39*(8), 591-598.

Ladd, H., & Lauen, D. (2010). Status versus growth: The distributional effects of school accountability policies. *Journal of Policy Analysis and Management, 29*(3), 426-450.

Ladd, H. (2012). Presidential address: Education and poverty: Confronting the evidence. *Journal of Policy Analysis and Management, 31*(2), 203-227.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis, 24*(1), 37-62.

Lleras-Muney, A. (2005). The relationship between education and adult mortality in the United States. *Review of Economic Studies, 72*(1), 189-221.

Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review, 94*(1), 155-189.

Mischel, W., Shoda, Y., & Rodriguez, M. (1989). Delay of gratification in children. *Science, 244*(4907), 933-938.

National Center for Education Statistics. (2012). Mathematics 2011: National Assessment of Education Progress at Grades 4 and 8. *U.S. Department of Education: NCES 2012-458.* Available from: http://nces.ed.gov/nationsreportcard/pdf/main2011/2012458.pdf.

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

Psacharopoulos, G., & Patrinos, H.A. (2004). Returns to investment in education: A further update. *Education Economics, 12*(2), 111-134.

Reininger, M. (2012). Hometown disadvantage? It depends on where you're from: Teachers' location preferences and the implications for staffing schools. *Educational Evaluation and Policy Analysis, 34*(2), 127-145.

Robinson, V., Lloyd, C., & Rowe, K. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly, 44*(5), 635-674.

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

Rockoff, J., Jacob, B., Kane, T., & Staiger, D. (2011). Can you recognize an effective teacher when your recruit one? *Education Finance and Policy, 6*(1), 43-74.

Ross, C., & Wu, C. (1995). The links between education and health. *American Sociological Review, 60*(5), 719-745.

Rouse, C. (2005). The labor market consequences of an inadequate education. Prepared for the Equity Symposium on *The Social Costs of Inadequate Education* at Teachers' College, Columbia University. Available from: www.literacycooperative.org/documents/TheLaborMarketConsequencesofanInadequateEd.pdf.

Sewell, W., & Hauser, R. (1972). Causes and consequences of higher education: Models of the status attainment process. *American Journal of Agricultural Economics, 54*(2), 851-861.

Thompson, C.L., Brown, K.M., Townsend, L.W., Henry, G.T., & Fortner, C.K. (2011). Turning around North Carolina's lowest achieving schools (2006-2010). *Consortium for Educational Research and Evaluation—North Carolina.* Available from: http://publicpolicy.unc.edu/research/publications-presentations-and-reports.

Walker, M. (2011). PISA 2009 plus results: Performance of 15 year-olds in reading, mathematics, and science for 10 additional participants. *Australian Council for Education Research.* Available from: https://mypisa.acer.edu.au.

CHAPTER ONE

Do Teachers' Non-Cognitive Skills and Traits Predict Effectiveness and Evaluation Ratings?[1]

SUMMARY

Building upon research in economics and psychology that shows the significant effects of individuals' non-cognitive characteristics on outcomes of interest, this paper uses data from Teach For America's corps member selection process to ask whether measures of teachers' non-cognitive characteristics predict value-added and evaluation ratings of teacher competencies. Overall, results indicate that non-cognitive skills and traits exert a significant influence on teacher quality. Teachers' organizational ability and academic/behavioral expectations of students predicted value-added gains in elementary grades and high school, respectively, while expected non-cognitive characteristics (e.g. leadership) predicted higher ratings for specific teacher competencies (e.g. teachers demonstrate leadership). This research represents a "proof of concept" concerning the influence of non-cognitive characteristics; the importance of this work for policy—better equipping preparation programs and school districts to recruit and select/hire teachers and cultivate these traits—necessitates continued research on a more representative sample of teachers.

[1] Kevin C. Bastian is the sole author on this paper and completed all data management, analysis, and writing tasks. Gary T. Henry provided editing comments and feedback regarding the framing of the paper and the interpretation of results.

INTRODUCTION

Increasingly, a rich body of research in economics and psychology indicates that individuals' non-cognitive skills and traits—e.g. perseverance, self-control, motivational ability—significantly influence academic and labor market outcomes of interest. For example, Duckworth and Seligman show that measures of self-discipline, not IQ scores, more accurately predict middle school students' grade point averages (Duckworth & Seligman, 2005), while Heckman and Rubinstein find that differences in non-cognitive characteristics explain the wage differentials between high school graduates and GED holders (Heckman & Rubinstein, 2001). Further evidence suggests that individuals' non-cognitive characteristics are malleable and can be cultivated through direct intervention (Mischel, Shoda, & Rodriguez, 1989).

Given (1) the sizable effects of teachers on both short-term student achievement and longer-term outcomes, such as collegiate attendance and job earnings (Aaronson, Barrow, & Sander, 2007; Chetty, Friedman, & Rockoff, 2011; Hanushek, 2011; Nye Konstantopoulos, & Hedges, 2004; Rockoff, 2004); and (2) the limited amount of variation in teacher effectiveness explained by teacher professional credentials (Boyd, Goldhaber, Lankford, & Wyckoff, 2007; Clotfelter, Ladd, & Vigdor, 2007, 2010; Greenwald, Hedges, & Laine, 1996; Staiger & Rockoff, 2010), a key question for education policy research is whether non-cognitive skills and traits predict teacher outcomes of interest. Here, anecdotal evidence suggests positive effects. For instance, facing the rigors of high-need schools, teachers' levels of perseverance—working hard to overcome adversity—may influence effectiveness and persistence in the profession (Haberman, 1995; Stotko, Ingram, & Beaty-O'Ferrall, 2007). Likewise, recognizing that

classroom learning is dependent upon the cooperation and investment of students, a teacher's motivational ability—making students believe they can and want to learn course material—may impact student achievement gains (Ames, 1990; Cohen, 2011).

As of yet, however, rigorous quantitative evidence regarding the effects of teachers' non-cognitive characteristics is limited. Therefore, to better understand why teachers succeed, I employ a unique dataset of eight pre-service teacher traits, many of them non-cognitive in nature, measured by Teach For America during its corps member selection process and ask the following questions:

1) Do pre-service measures of teachers' non-cognitive skills and traits predict value-added effectiveness?
2) Do pre-service measures of teachers' non-cognitive skills and traits predict evaluation ratings of teacher competencies?

Overall, the results of this work suggest that non-cognitive skills and traits influence teacher performance as measured by value-added models and principals' evaluation ratings. In elementary grades (mathematics and reading) teachers' organizational ability most strongly predicted students' achievement gains, while in high school, teachers' respect (academic and behavioral expectations) for students significantly predicted value-added effectiveness. Concerning evaluation ratings, two non-cognitive characteristics, leadership and motivational ability, predicted higher scores for multiple teacher competencies. Furthermore, expected non-cognitive characteristics (e.g. leadership) predicted higher ratings for specific teacher competencies (e.g. teachers demonstrate leadership).

While these findings illustrate the effects of non-cognitive traits, the highly-select research sample—TFA corps members—necessitates further work to determine whether and how the effects of non-cognitive skills and traits generalize to a wider population of teachers. Therefore, this research represents a "proof of concept" and impetus for researcher-practitioner partnerships to gather similar non-cognitive data and test effects. The policy implications of such work are clear, as findings could better equip teacher preparation programs and school districts to: (1) recruit and select/hire prospective teachers and (2) actively develop or cultivate non-cognitive skills and traits.

In the following sections I first present research evidence regarding the significant, yet limited effects of teachers' cognitive ability and then provide background and review evidence regarding non-cognitive skills and traits. Next, I detail the data and sample, particularly the eight traits measured by TFA. Then, I describe the analytical plan for addressing each research question. Finally, I present the results from the analyses of teacher effectiveness and evaluations of teacher competencies and conclude with a discussion of research and policy significance.

BACKGROUND

The Effects of Cognitive Ability on Teacher Effectiveness

Given the large effects of teachers on student achievement there is considerable interest from researchers and policymakers to identify teacher credentials that predict effectiveness. One often-studied characteristic are indicators of teachers' cognitive ability, such as SAT/ACT scores, selectivity of the undergraduate institution attended, or licensure exam scores. Here, longitudinal data show that when measured by IQ scores or campus selectivity the cognitive ability of the teacher workforce has declined over the

last several decades, largely in response to greater labor market opportunities for women (Bacolod, 2007; Corcoran, 2007).

However, research evinces that greater levels of cognitive ability exert small, but significant effects on student achievement. Findings from the *Coleman Report* indicated that teacher verbal ability was significantly associated with student achievement, and despite issues with data aggregation to the school level, early meta-analyses returned positive effects for teacher cognitive ability (Coleman, Campbell, Hobson, McPartland, Mood, Weinfeld, et al., 1966; Greenwald, Hedges, & Laine, 1996; Wayne & Youngs, 2003). More recent analyses with teachers directly linked to students have shown that teachers: (1) graduating from highly competitive undergraduate institutions generate more student test score growth than peers from less competitive universities (Clotfelter, Ladd, & Vigdor, 2006); (2) scoring higher on all exams or on licensure-specific content and pedagogical exams (e.g. Praxis tests) are more effective than lower scoring peers—this effect is stronger at the tails of the test score distribution (Clotfelter, Ladd, & Vigdor, 2007, 2010; Goldhaber, 2007); and (3) rating higher in a latent cognitive ability index produce larger student achievement gains than peers with a lower index score (Rockoff, Jacob, Kane, & Staiger, 2011). Furthermore, analyses of academically competitive teacher selection programs in North Carolina and New York indicate that (1) program graduates are more effective than other novice teachers due to selection on cognitive ability (Henry, Bastian, & Smith, 2012) and (2) individuals who receive higher admissions rankings by the teacher selection program generate larger student achievement gains than peers who receive lower rankings (Rockoff & Speroni, 2011).

Overall, the research evidence suggests that recruiting and selecting/hiring prospective teachers with higher levels of cognitive ability may modestly improve the effectiveness of the teacher workforce. The small portion of variance in teacher effectiveness explained by such credentials, however, helps motivate a greater focus on non-cognitive skills and traits.

What are Non-Cognitive Characteristics?

Broadly defined, non-cognitive characteristics are those academically and occupationally relevant skills and traits that, while not specifically intellectual or analytical in nature, influence behavior and facilitate achievement (Rosen, Glennie, Dalton, Leonnon, & Bozick, 2010). Examples of such attributes include perseverance, motivation, and self-control. Importantly, the term non-cognitive should not imply a complete absence of cognitive ability, as cognition is present in almost all aspects of human behavior. Rather, cognition can be conceptualized across a spectrum, with certain skills/abilities more readily identified as cognitive and other traits more readily identified as non-cognitive. Despite this lack of a sharp contrast, cognitive ability can be conceptually and empirically separated from non-cognitive traits (Borghans, Duckworth, Heckman, and ter Weel, 2008).

The research disciplines most focused on non-cognitive characteristics are psychology and economics. In psychology, non-cognitive attributes are often referred to as personality traits, and most prominent among these are the "Big Five." Specifically, the Big Five personality traits, openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism, are the five factors that consistently emerge from personality assessments (Digman, 1990). These factors represent personality at the

highest level, with each factor summarizing a host of distinct, specific personality characteristics. For instance, within the conscientiousness factor are personality facets such as order, dutifulness, competence, and self-discipline. Issues with the Big Five include: (1) the exclusion of certain traits, especially individual motivation and (2) controversy concerning to which factor certain personality facets belong (Borghans, Duckworth, Heckman, & ter Weel, 2008). While both psychologists and economists study the relationship between non-cognitive characteristics and later outcomes of interest—academic and occupational—studies in economics lack the unifying structure of the Big Five. Instead, economists tend to examine a single or small group of traits, focusing on effects and their implications for policy—for example, Heckman has focused on the Perry Preschool Program in studies assessing the cost-effectiveness of policymakers investing in programs that develop the non-cognitive characteristics of children (Heckman, 2006). Overall, the present study is more aligned with this economics tradition, asking how these eight teacher traits influence teacher quality and what the findings mean for research and policy.

The Effects of Non-Cognitive Skills and Traits

Beyond cognitive ability, researchers in economics and psychology have long been interested in the relationships between outcomes of interest (academic achievement, labor market success) and individuals' non-cognitive skills and traits. Overall, a rich body of research evidence clearly supports both the significance and malleability of these characteristics. Below, I highlight the results of a few noteworthy studies.

Regarding the effects of non-cognitive skills and traits on students' academic achievement, research finds that: (1) the cognitive ability of low-IQ students is not fixed,

but rather, increases significantly with motivation—the introduction of a performance incentive (candy) (Clingman & Fowler, 1976); (2) measures of self-control, indicated by whether a four year-old child is able to delay gratification and not consume a marshmallow, significantly predict higher SAT scores and enrollment at higher ranked colleges (Mischel, Shoda, & Rodriguez, 1989); and (3) measures of grittiness and perseverance explain success in National Spelling Bee competitions (Duckworth, Kirby, Tsukayama, Berstein, & Ericsson, 2011). Non-cognitive characteristics continue to influence outcomes of interest for adults, as research also shows that: (1) non-cognitive measures, not cognitive ability, explain the wage differentials between high school graduates and GED holders (Heckman & Rubinstein, 2001); (2) in the workplace, conscientiousness, one of the Big Five personality traits, consistently predicts job proficiency for five different occupational groups (Barrick & Mount, 1991); and (3) measures of childhood self-control significantly predict physical health, substance dependence, personal finances, and criminal offense outcomes at age 32 (Moffitt et al., 2011). Finally, research evidence indicates that these non-cognitive traits are not fixed, but rather, can respond to intervention as: (1) children coached in self-control strategies were better able to delay gratification in the marshmallow experiment (by 7-18 minutes depending on strategy) (Mischel, Shoda, & Rodriguez, 1989) and (2) high-school students participating in mental contrasting activities—concentrating on both a positive outcome and steps needed to overcome the obstacles in the way—more diligently prepared for the SAT exam (Duckworth, Grant, Loew, Oettingen, & Gollwitzer, 2011).

Turning to teaching, efforts to measure the non-cognitive skills and traits of prospective teachers is not new. For example, many school districts employ scripted

questionnaires or interview protocols, such as the Haberman Star Teacher Pre-Screener or the Teacher Perceiver Interview, to identify prospective teachers' non-cognitive traits and make hiring decisions. Likewise, many teacher preparation programs measure and cultivate the dispositions of pre-service teachers. A primary reason that districts and teacher education programs engage in such activities is the belief that these traits directly influence the quality of teacher practices, and then indirectly, impact student academic outcomes. For instance: (1) higher levels of perseverance may cause teachers to work harder and try multiple instructional strategies to make an academic concept clear to a struggling student or (2) greater motivational ability may enable teachers to first invest their students in the academic goals of the class and then push students to exert the effort required to reach those goals.

With the increasing prevalence of student achievement data, a nascent research agenda is now able to investigate whether non-cognitive skills and traits produce the hypothesized effects on teachers' effectiveness. Here, initial studies have returned promising findings. For example, Rockoff and colleagues find that a one standard deviation increase in a latent non-cognitive construct—extraversion, conscientiousness, personal efficacy, general efficacy, and the Haberman Star Index total score—led to a significant improvement of 0.033 standard deviations in student math achievement and of 0.272 standard deviations in teachers' subjective evaluations (Rockoff, Jacob, Kane, & Staiger, 2011). In work with TFA corps members, Duckworth et al. show that a one standard deviation increase in measures of teachers' grittiness and life satisfaction are associated with a twenty-three and thirty-six percent increased probability, respectively, of students making significant academic gains (Duckworth, Quinn, & Seligman, 2009).

While this work represents an important advance, the results should be interpreted cautiously, as the measure of academic gains used did not come from state or district criterion-referenced exams, but rather, was teacher reported. In research focused on New York City TFA corps members, Dobbie finds that a one standard deviation increase in corps members' leadership and perseverance predict achievement gains in math of 0.054 and 0.040 standard deviations, respectively (Dobbie, 2011). In comparison to Dobbie's analysis, an advantage of the present study is a more comprehensive examination of teacher quality—both teacher value-added and evaluation ratings—and a more heterogeneous sample of school environments—schools in both urban and rural communities across North Carolina.

Overall, evidence suggests that non-cognitive characteristics may help explain what makes teachers effective, and longer-term, facilitate results-aligned selection/hiring practices for school districts and preparation programs. The present study contributes to this nascent research agenda by examining the effects of non-cognitive skills and traits on two teacher quality outcomes—value-added effectiveness and evaluation ratings of teacher competencies. Below, I provide background information regarding Teach For America and describe the eight pre-teaching measures that are the focus of this work.

DATA AND SAMPLE

Background on Teach For America

Founded in 1990, Teach For America is an organization committed to closing the achievement gap between high and low-income students by recruiting academically competitive recent college graduates and placing them into low-income schools and communities to teach. Because most corps members do not enter the program with a

background in teacher education, TFA provides both an intensive, five-week Summer Institute prior to corps members' first year teaching and on-going coaching and professional development. Corps members commit to teach in their low-income placement schools for at least two years; evidence from administrative data suggests that a large majority of corps members fulfill this two-year commitment but only a minority continue teaching beyond this point (Henry, Bastian, & Smith, 2012).[2] Existing research on TFA corps members' effects on student achievement gains returns mixed results (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; Darling-Hammond, Holtzman, Gatlin, & Heilig, 2005; Decker, Mayer, & Glazerman, 2006; Henry, Bastian, Fortner, Kershaw, Purtell, Thompson, & Zulli, 2013; Kane, Rockoff, & Staiger, 2008; Raymond, Fletcher, & Luque, 2001; Xu, Hannaway, & Taylor, 2011). Over time, however, the trend in evidence suggests that corps members are effective at promoting student achievement growth, especially in STEM courses (mathematics and science) and at the secondary school level (Boyd, Grossman, Hammerness, Lankford, Loeb, Ronfeldt, & Wyckoff, 2012; Decker, Mayer, & Glazerman, 2006; Henry, Bastian, & Smith, 2012; Henry et al., 2013; Xu, Hannaway, & Taylor, 2011).

Most relevant to the present study is the process by which TFA selects applicants for admission into the program. Essentially, TFA collects data on applicants' cognitive ability and non-cognitive characteristics throughout its application process—submission of documents, such as a letter of intent, resume, college transcript, and letters of recommendation; a phone interview; and an in-person interview—and based upon internal analyses showing how these traits predict corps member effectiveness, TFA uses

[2] Evidence from three cohorts of TFA corps members in North Carolina (2004-05, 2005-06, and 2006-07) shows that approximately 90% of corps members fulfill their two year teaching commitments, but only 30% of corps members continued into a third year of teaching (Henry, Bastian, & Smith, 2012).

this applicant information to make data-driven admissions decisions (Ripley, 2010). Below, I detail the measures of cognitive ability and non-cognitive traits measured by TFA.

Teach For America Data

To examine the effects of non-cognitive skills and traits on teachers' effectiveness and ratings of teacher competencies, I employ a unique dataset of eight pre-teaching measures collected by TFA and used during its corps member selection process. Two of the selection criteria are more cognitive in nature—prior academic achievement and critical thinking skills—while the remaining six criteria are more non-cognitive in nature—leadership, perseverance, organizational ability, motivational ability, respect for low-income students and communities, and fit with TFA's mission to close the achievement gap (See the top portion of Table 1.1 for the mean and standard deviation of each trait). Below, I briefly detail what each criterion measures and review the theoretical and empirical relation between each trait and teaching quality. Here, it should be noted that due to the level of selectivity of TFA corps members, which I further detail in the research sample section, results for these measures may differ from results for a more generalizable sample of teachers.[3] I conclude this section with the results of a factor analysis to determine whether TFA measures eight unique traits or a smaller number of latent constructs.

Prior academic achievement: This criterion captures whether an individual has achieved ambitious, measurable results in prior academic work. Empirical research

[3] Prior research evinces that, on average, TFA corps members have significantly higher levels of cognitive ability than other novice teachers (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006). Nascent evidence collected by Duckworth and colleagues also suggests that corps members may have significantly higher levels of non-cognitive characteristics (Duckworth, Quinn, & Seligman, 2009).

indicates that teachers with higher levels of cognitive ability—SAT/ACT scores, college GPA, licensure exams (Praxis), or rankings within teacher preparation programs—produce larger student test score gains than lower-scoring peers (Goldhaber, 2007; Greenwald, Hedges, & Laine, 1996; Kukla-Acevedo, 2009; Rockoff, Jacob, Kane, & Stagier, 2011; Rockoff & Speroni, 2011).

Critical thinking skills: This criterion captures an individual's ability to accurately assess the cause of problems and generate effective solutions. While little quantitative work examines whether this teacher trait predicts student achievement gains, many rubric-based teacher observation and evaluation protocols specifically rate teachers' competency to reflect on practice—requiring critical thinking skills—and implement changes to address problems or shortcomings (Danielson, 2007; McREL, 2009). As data from these protocols becomes available, researchers may be able to determine whether critical thinking predicts a teacher's ability to promote student achievement gains.

Leadership: This measure evaluates an individual's performance and experience leading others in jobs or extracurricular activities. A considerable amount of research focuses on the relationship between leadership opportunities for teachers and school outcomes, yet little work connects measures of individual teachers' leadership experience to student achievement. Here, the conceptual model is that high-quality teachers are comparable to leaders in other professions and settings—effective leaders set goals, invest others (students) in those goals, create plans to reach goals, execute, and reflect on those plans until fruition (Farr, 2010).

Perseverance: This measures the extent to which an individual, when challenged, works purposefully and relentlessly to achieve goals. Education researchers have long identified perseverance—the ability to overcome obstacles—as a key characteristic of successful teachers, particularly in challenging school and classroom environments (Haberman, 1995; Stotko, Ingram, & Beaty-O'Ferrall, 2007). Recent quantitative research focused on student test score outcomes corroborates this finding, showing that teachers with higher levels of grittiness or perseverance produce larger student achievement gains than peers with lower perseverance scores (Dobbie, 2011; Duckworth, Quinn, Seligman, 2009).

Organizational ability: This measures the extent to which an individual plans well—keeping the goal in mind—and effectively manages tasks until completion. Outside education, work in psychology on the Big Five personality traits indicates that conscientiousness—related to organizational ability—consistently predicts job performance (Barrick & Mount, 1991; Judge, Higgins, Thoreson, & Barrick, 1999). Education researchers also connect organization and planning to teaching success, especially given the recent focus on backwards planning—planning with the final objective in mind—as a key component of reaching academic goals (Haberman, 1995; Wiggins & McTighe, 2005).

Motivational ability: This criterion evaluates an individual's ability to use interpersonal skills to motivate and lead others toward a common goal. While few quantitative studies examine the link between this teacher trait and student achievement gains, considerable research work asserts that effective teaching requires investing

students and helping them believe that they can and want to reach academic goals (Ames, 1990; Cohen, 2011).

Respect for low-income students and communities: This indicates the extent to which an individual holds high academic and behavioral expectations for low-income and low-achieving students and communities. As defined, respect closely relates to locus of control—a teacher's belief that explanations for student academic success reside in the actions of teachers and students themselves (Dembo & Gibson, 1985; Haberman, 1995; Murray & Staebler, 1974). Considerable research indicates that high expectations can be predictive of student achievement gains (Brophy, 1983; Brophy, 1986; McKown & Weinstein, 2008; Rosenthal & Jacobsen, 1966).

Fit with the TFA mission: This measures an individual's understanding of and commitment to the TFA mission to close the achievement gap between students in high-poverty versus low-poverty communities and schools. While this criterion, in comparison to those above, is more specific to TFA, its underlying construct—a primary focus on significantly improving student achievement—is supported as a characteristic of successful schooling and teaching, particularly in high-need schools (Haberman, 1995; Henry, Thompson, Brown, Cunningham, Kainz, Montrosse, Sgammato, & Pan, 2008; Thompson, Brown, Townsend, Henry, & Fortner, 2011).

Factor Analysis: To better understand the relationships among the pre-service teacher traits and determine whether TFA measures eight unique characteristics or fewer underlying constructs, I first examined the correlations among the traits. Perhaps surprisingly, the middle portion of Table 1.1 evinces that few of the traits exhibit strong relationships with each other. The largest correlation is between respect and fit with the

TFA mission at 0.27, while most of the remaining traits have correlations less than 0.10 in absolute value. This is a preliminary indication that these characteristics may represent unique traits of the prospective teachers.

To test these relationships more rigorously, I performed principal components factor analysis with an oblique rotation procedure. Using Eigenvalues greater than one as a cut-off to identify factors, this approach originally returned a three factor structure, with respect and fit with the TFA mission loading together, organizational ability and motivational ability loading together,[4] and achievement, leadership, perseverance, and critical thinking all loading together. While the first two factors strongly loaded in the same direction, the last factor had items strongly loading both positively (achievement and critical thinking) and negatively (leadership and perseverance). Upon considering this factor one clear distinction was that the two items loading positively—achievement and critical thinking—are cognitive in nature, while the two items loading negatively—leadership and perseverance—are non-cognitive in nature. This suggested a way to separate these items into two conceptually distinct factors. To examine this possibility further, I correlated teachers' Praxis II licensure exam scores, a proxy for cognitive ability, with the eight skills and traits measures. As shown at the bottom of Table 1.1, the Praxis II measure correlated positively with achievement and critical thinking at 0.271 and 0.186, respectively, while it correlated negatively or not at all with leadership and perseverance.[5]

[4] The groupings of these first two factors is not surprising, since respect and fit had the highest pair-wise correlation value of 0.272 and organization ability and motivational ability had the second highest pair-wise correlation value of 0.190.

[5] The middle portion of Table 1.1 also shows that prior academic achievement and critical thinking correlate positively with each other and negatively with the six remaining traits (considered non-cognitive).

Therefore, with both conceptual and empirical support, I created separate factors for prior achievement and critical thinking and leadership and perseverance. The final factor structure includes one factor that is cognitive in nature—prior academic achievement and critical thinking—and three factors that are non-cognitive in nature—leadership and perseverance, organizational ability and motivational ability, and respect and fit with the TFA mission. Due both to the weak correlations between the eight selection criteria (shown in Table 1.1) and the challenge of drawing research and policy implications from analyses in which factors (combinations of traits) are the focal variables, I table and discuss results for the eight individual traits in the findings section. I include results from models with the four factors in Appendix A.

Research Sample

The sample for this research consists of individuals accepted into TFA from 2007-08 through 2010-11 and placed into one of the organization's two North Carolina regions (Charlotte or Eastern North Carolina). As shown in Table 1.2, this sample is distinguished from other beginning teachers in two ways. First, TFA corps members entered the teaching profession with greater amounts of cognitive ability than other novice peers. For example, TFA corps members score one-third of a standard deviation higher on Praxis II licensure exams, and the percentage of corps members graduating from an undergraduate institution with a very, highly, or most competitive Barron's ranking is three times greater than other novice teachers. Second, in comparison to other novice tested-grade/subject teachers, TFA corps members work in particularly challenging school and classroom environments. For instance, corps members teaching

---

This is further support to separate prior academic achievement and critical thinking from leadership and perseverance.

in a tested-grade in elementary schools work in (1) schools in which students passed only 43 percent of the End-of-Grade exams taken and 87 percent of the student body qualifies for subsidized school lunches and (2) classrooms in which the average prior achievement score is one-half of a standard deviation below the statewide mean and 80 percent of students are eligible for subsidized school lunches. It is these individual and workplace characteristics of corps members that necessitate continued research on a more representative sample of teachers and environments. Below, I describe the specific sample of corps members and study years included in the analyses of teacher effectiveness and evaluation ratings of teacher competencies.

Teacher Effectiveness: To determine whether non-cognitive skills and traits predict teacher effectiveness, the research sample for this question includes all corps members from the full sample—selected into TFA from 2007-08 through 2010-11—teaching in a tested grade or subject in North Carolina public schools. In elementary (3-5) and middle grades (6-8) this includes mathematics and reading, while at the high school level (grades 9-12), this includes ten End-of-Course (EOC) exams—algebra 1, algebra 2, geometry, biology, physical science, chemistry, physics (STEM subjects), English 1, U.S. history, and civics/economics (non-STEM subjects). Overall, I ran two value-added models at the elementary grades level (mathematics and reading), two value-added models at the middle grades level (mathematics and reading), and three value-added models at the high school level (all ten EOC exams combined, STEM subjects only, and non-STEM subjects only). Furthermore, for each of these subject or grade-specific analyses I also limited models to: (1) first year teachers only, to determine whether non-cognitive skills and traits have immediate effects on teacher effectiveness;

(2) second year teachers only, to determine whether the effects of non-cognitive skills and traits develop or dissipate with experience; and (3) first and second year teachers combined, to determine whether non-cognitive skills and traits have an overall effect on student test score gains. Because TFA has a two-year teaching commitment and a majority of corps members in North Carolina public schools do not persist into a third year, I did not perform value-added analyses with more experienced teachers.

Evaluation Ratings: In the 2010-11 school year North Carolina piloted the McREL teacher evaluation system, a rubric-based observation and evaluation protocol with which principals rate teacher competencies across five standards (detailed in the dependent variables section below). To examine whether non-cognitive skills and traits predict teachers' evaluation ratings, the research sample for this question consists of all corps members from the full sample—selected into TFA from 2007-08 through 2010-11—rated by their school principal during the 2010-11 school year (251 teachers total; 249 with non-missing covariates for analyses). Due to the small size of this sample, I do not perform any sub-analyses by teacher assignment type (grade level, tested vs. non-tested grade/subject); future analyses, with larger samples, may benefit from separately investigating the effects of non-cognitive skills and traits in these areas.

Dependent Variables and Covariates

Dependent Variables: The dependent variable for the teacher effectiveness analyses is students' current test score performance on the North Carolina End-of-Grade (grades 3-8) mathematics and reading exams—standardized within subject, grade, and year—or the End-of-Course high school exams—standardized within subject (e.g. algebra 1) and year. Standardized mathematics and reading scores from the previous

grade, or from 8th grade for high school students, serve as the measure of prior achievement in these value-added models.[6] More on the control variables included to isolate the effects of non-cognitive skills and traits in the teacher effectiveness analyses is included in the covariates section below.

For analyses of teacher competency ratings, the dependent variable comes from the McREL teacher evaluation protocol. The North Carolina State Board of Education approved this new evaluation rubric in October 2008 and schools piloted the evaluation system in the 2010-11 school year. McREL contains five standards on which principals rate teacher competency: (1) teachers demonstrate leadership (2) teachers establish a respectful environment for a diverse group of students (3) teachers know the content they teach (4) teachers facilitate learning for their students and (5) teachers reflect on their practice. Within each of these standards, principals document the presence of key teacher behaviors. For instance, with standard five—teachers reflect on practice—principals focus on whether teachers collect student assessment data to analyze learning, adapt practice based on data, and participate in professional development to address areas for professional growth. To evaluate probationary teachers (the sample of corps members for this work), principals conduct at least three formal observations during the school year. Prior to the first observation principals and teachers have a pre-conference and after each formal observation, principals and teachers have a post-conference. Finally, at the end of the school year, principals and teachers have a summary evaluation conference and for each of the five McREL standards principals officially rate teachers as either: not demonstrated, developing, proficient, accomplished, or distinguished. Given prior

[6] In elementary grades students' prior achievement scores are an average of the standardized mathematics and reading performance. Furthermore, I include 3rd grade students in value-added analyses in 2007-08 and 2008-09 by using the EOG mathematics and reading pre-tests given at the beginning of 3rd grade.

research indicating that many teacher evaluation systems do not distinguish effectively between teachers of differing quality, I created outcome measures from the McREL ratings that identify a select sample of teachers (Toch & Rothman, 2008; Weisberg, Sexton, Mulhern, & Keeling, 2009). Specifically, across each of the five evaluation standards 'proficient' was the modal scoring category for corps members, with principals rating approximately 70 to 80 percent of corps members at 'proficient' or below.[7] Therefore, for each of the five evaluation standards I created a dichotomous dependent variable for whether or not the principal rated the teacher 'above proficient'—either 'accomplished' or 'distinguished.' This specification allowed me to determine whether: (1) non-cognitive skills and traits predict higher ratings of teacher competency and (2) different non-cognitive skills and traits predict different facets of teachers' practices (See Appendix Table A1 for descriptive information on corps members' ratings and the ratings of all other teachers evaluated in the 2010-11 school year).

Covariates: The focal variables for the analyses of teacher effectiveness and teacher competency ratings are the eight individual traits—six of them non-cognitive in nature—measured by TFA during its selection process. All results for models with the four factors are included in Appendix A. For analyses, I standardized each of the traits within cohort-year and entered the eight standardized traits into models, collectively, to identify their independent effects on teacher effectiveness or ratings of teacher competencies. As detailed in Table 1.3, to further isolate the effects of these non-cognitive skills and traits on students' adjusted average test score growth, I included a rich set of student, classroom, teacher (including experience controls for models with 1st

[7] There were no TFA corps members in my sample rated as 'not demonstrated' for any of the five McREL standards and only two to five percent of corps members were rated as 'developing' for any of the five McREL standards.

and 2nd year corps members), and school covariates in value-added models. Finally, in logistic regression analyses predicting whether principals rate teachers above proficient for the five evaluation standards, I controlled for teacher experience—since more experienced teachers are significantly more likely to be highly rated—and the same set of school covariates as listed in Table 1.3.

ANALYSIS PLAN

Teacher Effectiveness

The goal of this analysis was to isolate the effects of teachers' non-cognitive skills and traits on their students' adjusted-average test score gains. To do this I utilized the extensive set of administrative data provided by the North Carolina Department of Public Instruction to estimate a value-added model with a rich set of student, classroom, teacher, and school covariates (Table 1.3). Here again, I ran separate models for first year corps members, second year corps members, and a combined model of first and second year corps members to examine how the effects of non-cognitive skills and traits may change early in teachers' careers. I also used cluster-adjusted standard errors at the teacher level to account for the clustering of students within teachers that, if left unadjusted, could result in reduced standard errors and false positive hypothesis tests.[8] The equation used to estimate the effects of the individual teacher traits on corps member effectiveness is as follows:

$$A_{ijst} = \alpha A_{it-n} + \beta Traits_j + \gamma X_{ijst} + \delta C_{jst} + \theta S_{st} + \varepsilon_{ijst} \quad (1)$$

where A_{ijst} is the test score for student *i* taught by teacher *j* in school *s* at time *t*;

[8] I chose to cluster standard errors at the teacher level for two reasons: (1) in these models my sample is limited to TFA corps members only and the non-cognitive characteristics of these teachers represent the "treatment" of interest and (2) clustering at the teacher level follows the procedures set forth by Dobbie (2011) with similar TFA selection data in New York City.

A_{it-n} represents the prior test score(s) for student *i*;

$Traits_j$ represents a vector of the eight standardized individual teacher traits;

β estimates the average effect of the eight individual teacher traits on students' adjusted-average test score growth;

X_{ijst} represents a set of time-invariant and varying individual student characteristics;

C_{jst} represents a set of classroom and teacher characteristics;

S_{st} represents a set of school characteristics;

and ε_{ijst} is a disturbance term representing all unexplained variation in student achievement.

In response to the non-random assignment of teachers to students and the influence of unmeasured school contextual factors on teacher effectiveness, many value-added estimation approaches also include a series of specification checks, such as student fixed effects or school fixed effects, to mitigate internal validity threats (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; Clotfelter, Ladd, & Vigdor, 2005; Henry, Bastian, & Smith, 2012; Goldhaber & Liddle, 2011; Lankford, Loeb, & Wyckoff, 2002; Rothstein, 2010). However, due to the nature of this analysis—examining the effects of non-cognitive skills and traits within the select sample of TFA corps members—typical econometric approaches (fixed effects) to address endogeneity concerns are both: (1) not practical, since few students experience corps members in consecutive years in the same subject (student fixed effects) and few corps members teach the same tested subject within the same schools (school fixed effects) and (2) likely unnecessary, since corps members generally teach in similar school and classroom environments with many low-income and low-achieving students. This classroom and school placement mitigates the

primary endogeneity threat—that within the sample of corps members, individual teacher traits have positive associations with assignment to students. As comparable datasets become available for a larger, more generalizable sample of teachers, these fixed effect estimation approaches may be feasible to better isolate the effects of non-cognitive skills and traits.

Finally, to provide a broader context for the non-cognitive skills and traits results, given the select estimation sample of TFA corps members, I include additional value-added models comparing the effectiveness of TFA corps members to all other novice teachers.[9] Specifically, I omit the individual teacher trait variables from equation one, insert an indicator variable for TFA, and specify models comparing: (1) first year corps members with all other first year teachers; (2) second year corps members with all other second year teachers; and (3) a combined model with first and second year teachers.[10] Results from these models provide a basis of comparison, or means to quantify the size of the non-cognitive skills and traits estimates from equation one and help illustrate that in this select sample even if the average effect of a trait is negative (positive), the overall effectiveness of the corps members may be positive (negative).

Evaluation Ratings of Teacher Competencies

Because (1) many important aspects of teaching quality, such as reflecting on practice, assuming school and departmental leadership roles, or establishing a respectful classroom environment, may not be well-captured by value-added outcomes; (2) only

[9] Because these are statewide value-added models I cluster standard errors at the school-by-year level—the highest level of relevant sorting in the analysis.

[10] Since these models have a much larger sample size I include specification checks with school fixed effects. These fixed effects results are comparable to the results from the main models and are available upon request.

35% of classroom teachers teach in tested grades/subjects; and (3) improvements in teachers' instructional quality are generally a pre-cursor to student achievement gains, I also examined whether non-cognitive skills and traits predict evaluation ratings of teacher competencies. For these models I specified a dichotomous dependent variable—rated above proficient—for each of the five evaluation standards and used a logistic regression framework, controlling for teacher experience and a set of school contextual factors, to estimate odds ratios for the relationship between teacher traits and the evaluation score. Here, I included cluster-adjusted standard errors—at the school level—to account for dependence in the data. Overall, the equation to estimate the effects of the individual teacher traits is as follows:

$$\Pr(Above\_proficient_{js} = 1) = \frac{\exp(Traits_j + Exp_j + School_j)}{1 + \exp(Traits_j + Exp_j + School_j)} \quad (2)$$

where $Above\_proficient_{js}$ is a binary outcome equal to 1 for teacher *j* and evaluation standard *s* if the school principal rated the teacher above the 'proficient' category;

$Traits_j$ represents a vector of the eight standardized individual teacher traits;

Exp_j represents a set of single-year teacher experience indicators, in reference to first year corps members;

and $School_j$ represents a vector of school contextual factors.

Although principals received standard training on how to use the McREL teacher evaluation protocol prior to its roll-out in the 2010-11 school year, it is possible that, across principals, differences exist in ratings of teachers' competencies. Ordinarily, a school fixed effect added to equation two would mitigate this concern—limiting comparisons to teachers a single principal evaluates—but as with the value-added

models, the small sample of teachers with evaluation ratings working in the same schools precludes such an estimation approach. Therefore, while the findings are suggestive of relationships between teacher traits and evaluation ratings, these models cannot rule out rating tendencies of individual principals.

FINDINGS

Teacher Effectiveness

The goal of the first research question was to determine whether individual teacher traits, particularly those that are non-cognitive in nature, predict teacher effectiveness. As detailed in Table 1.2, these estimates of interest make comparisons within a highly-select sample of teachers and school/classroom environments, meaning value-added results may generalize differently to a full sample of instructors. Therefore, this work represents a "proof of concept" and opportunity for hypothesis generation regarding the effects of non-cognitive skills and traits. Below, I present teacher effectiveness findings at the elementary school and high school levels.[11]

Examining Table 1.4, non-cognitive skills and traits, particularly teachers' organizational ability, exhibit strong relationships with teacher value-added in elementary grades.[12] In mathematics models the effect of organizational ability is significant for second year teachers, while in reading, organizational ability significantly predicts effectiveness across all three models. These results (1) are consistent with prior research findings from psychology which indicate that conscientiousness—related to

[11] Value-added results in middle grades mathematics and reading models revealed few significant relationships between individual teacher traits and teacher effectiveness. Therefore, I focus on elementary and high school findings. Middle grades results are available upon request from the author.

[12] Most elementary grades classrooms in North Carolina are self-contained, meaning teachers will have value-added outcomes in both mathematics and reading for the same group of students.

organizational ability—predicts workplace performance (Barrick & Mount, 1991) and (2) suggest that teachers' ability to plan and manage tasks is important in elementary grades. Additionally, teachers' motivational ability, which loaded onto the same factor as organizational ability, significantly predicts second year teacher effectiveness in both mathematics and reading. In contrast to these positive findings, increased levels of teacher perseverance predict a significant decrease in achievement for students taught by second year teachers. This result is particularly unexpected since higher levels of perseverance were anticipated to help teachers succeed, especially in low-performing environments. Finally, there is one significant relationship between a cognitive trait and teacher effectiveness in elementary grades models—prior academic achievement predicts value-added for first year reading teachers.

Turning to high school (Table 1.5), results for non-cognitive characteristics indicate that teachers' respect—holding high academic and behavioral expectations for students—is significant in all three models for the ten high school End-of-Course exams, for second year STEM teachers, and in the second year and combined model for non-STEM teachers. This suggests that locus of control—a teacher's belief that explanations for student academic success reside in the actions of teachers and students—may be an important characteristic of high school teachers. While the effects of teachers' organizational and motivational ability are not as frequently significant in high school as in elementary grades, there are positive effects for second year high school teachers for both traits. This suggests that a possible direction for future research is to determine whether teachers' organizational and motivational ability is a reliable predictor of effectiveness across grade levels. As with the elementary grades results, perseverance

has a negative effect on the achievement of students taught by second year STEM teachers; even more consistently negative is the relationship between teachers' leadership ability and student achievement gains in non-STEM courses. Finally, value-added estimates indicate mixed results for cognitive traits. Across all high school End-of-Course exams and in models limited to STEM subjects, prior academic achievement positively predicts effectiveness for second year teachers, however, critical thinking negatively predicts effectiveness for the same sample.

As a way to quantify the magnitude of these effects, the bottom portions of Tables 1.4 and 1.5 present results comparing TFA corps members with all other first year, second year, and first and second year teachers. In elementary grades mathematics, for example, second year corps members outperform other second year teachers by 0.110 standard deviations, an effect equivalent to an additional 27 days of student learning in a 180 school-year. By comparison, a one-standard deviation increase in organizational ability for second year corps members produces student test score gains of 0.071 standard deviations—an effect equivalent to two-thirds of the overall TFA mathematics result and nearly 18 days of additional student learning.[13] In all high school EOC subjects a one standard deviation increase in teachers' respect is equivalent to one-fifth of the effect of TFA corps members versus all other novice instructors; in non-STEM subjects, the effect of respect is equivalent to one-half of the difference between corps members and other novice teachers. While these non-cognitive characteristics do not necessarily explain the comparative effectiveness of TFA corps members—the non-cognitive (or cognitive) characteristics explaining variability in TFA effects may be different than those driving

[13] Please see Henry, Thompson, Bastian, Fortner, & Marcus, 2011, for details concerning how to convert effectiveness estimates into equivalent days of student learning.

average differences between corps members and other novice teachers—these results do suggest that non-cognitive characteristics can have sizable and practically significant effects on student achievement.

Overall, these teacher effectiveness models returned mixed results concerning the effects of cognitive traits. In light of the high levels of cognitive ability for TFA corps members shown in Table 1.2—on average, corps members score one-third of a standard deviation higher on standardized licensure exams than other novice instructors—this finding should not suggest that teachers' cognitive ability is unimportant, but rather, that a more generalizable sample of teachers may be required to detect effects. Corroborating findings on non-cognitive characteristics from other research disciplines, value-added results indicated that teachers' non-cognitive skills and traits can significantly influence student achievement. More on these value-added findings, including the unexpected perseverance and leadership results and why significant effects were concentrated in teachers' second year, is included in the discussion section.

Evaluation Ratings of Teacher Competencies

To provide a more comprehensive view of teaching quality—especially important as many states and districts refocus their teacher observation and evaluation protocols to better distinguish between teachers—I examined whether individual traits predict a corps member earning a rating above proficient for the five standards of the North Carolina teacher evaluation rubric. Focusing on this broader measure of teaching quality is important for three reasons: (1) teachers engage in many actions that make positive contributions to schools and students that are not captured by value-added estimates; (2) only 35 percent of teachers teach in a tested grade/subject; and (3) improvements in

teachers' instructional practices, especially the classroom environment they create, their levels of content knowledge, and their ability to facilitate student learning, are likely a precursor to greater levels of student achievement.

Examining Table 1.6, results for standard one—teachers demonstrate leadership in the school and classroom—indicate that teachers' measured levels of leadership ability significantly increase the odds of a teacher receiving an evaluation rating above proficient. High ratings for standard two—teachers establish a respectful classroom environment—are predicted by academic achievement, leadership, motivational ability, and fit with the TFA mission. For the most cognitive of standards—teachers know the content they teach—one cognitive trait, prior academic achievement, and one non-cognitive trait, motivational ability, predict the odds of a teacher rating above proficient. Facilitating student learning, or teachers' ability to present ideas clearly, provide quality feedback to students, and check for understanding, is significantly predicted by academic achievement and leadership ability. Finally, higher levels of measured motivational ability significantly increase teachers' odds of being rated above proficient for the reflecting on practice standard.

Overall, three main findings emerge from these relationships between individual teacher traits and ratings of teacher competencies. First, non-cognitive characteristics, particularly leadership and motivational ability, exert a strong influence on teacher competency ratings. Second, although some results are consistent across analyses of teacher value-added and competency ratings (such as positive motivational ability effects for both outcomes of interest), there are differences in the significant relationships for value-added models and evaluation ratings. This suggests that principal ratings may

focus on different aspects of teaching quality than those related to student achievement gains. Finally, in several instances the relationship between specific traits and competency ratings was congruent with expectations. For example, teachers' measured leadership ability significantly predicted ratings above proficient on standard one (teachers demonstrate leadership). Further discussion regarding the implications of these findings is included below.

DISCUSSION

Given the significant effects of teachers on student academic outcomes, understanding what makes teachers effective is essential. Towards this end, prior research findings indicate that the professional credentials of teachers explain only a small portion of the variance in teacher effectiveness, and therefore, education policy is increasingly relying upon post-entry teacher performance data to make consequential human capital decisions (Gordon, Kane, & Staiger, 2006). Largely unexplored in this research agenda, however, is whether non-cognitive skills and traits, which significantly predict outcomes of interest in other research disciplines, influence aspects of teacher quality.

To address this gap, I employed a unique dataset of eight pre-service teacher traits measured by TFA during its corps member selection process. Overall, I considered two of the traits to be cognitive—prior academic achievement and critical thinking—and the remaining six traits to be non-cognitive—leadership, perseverance, organizational ability, motivational ability, respect, and fit with TFA. Factor analysis results supported this structure, separating into a single factor for cognitive ability and three factors for the non-cognitive characteristics. In value-added analyses, teachers' organizational ability most

strongly predicted teacher effectiveness in elementary grades, while in high school, teachers' respect for students significantly predicted achievement gains. These results connect with prior research from psychology showing the impact of conscientiousness on workplace performance and work in education indicating how locus of control and expectations influence student outcomes (Barrick & Mount, 1991; Dembo & Gibson, 1985; Rosenthal & Jacobsen, 1966). Additionally, higher levels of motivational ability predicted teacher value-added in elementary grades and high school overall. This result was not surprising, given that teacher success may be dependent upon the investment of students—making students believe they can and want to learn (Ames, 1990; Cohen, 2011). In comparison to the non-cognitive results, there were fewer significant findings for cognitive ability and the direction of these results was mixed—for instance, positive effects of prior academic achievement and negative effects of critical thinking in high school. Due to the selectivity of TFA, however, these results should not suggest that teachers' cognitive ability is unimportant, especially for a more representative sample of beginning teachers.

Overall, two of the value-added results warrant further discussion. First, there were few significant relationships between individual traits and first year teacher effectiveness, and instead, most significant value-added findings were concentrated in teachers' second year—for example, motivational ability in elementary grades or prior academic achievement in high school. One possible explanation for this finding is that the rigors of first year teaching limit the immediate effects of individual characteristics, but teachers with particular traits are able to more rapidly develop on-the-job between their first and second year of teaching. A question for future research, then, is whether

teachers' cognitive and non-cognitive characteristics explain initial differences in effectiveness or emerge as influential over a longer-term. Second, there were unexpected, negative results for both leadership and perseverance. Higher levels of leadership ability returned reduced student achievement gains in high school, while paradoxically, more measured perseverance, a trait which should benefit teachers, especially to overcome the challenges of high-need schools, adversely impacted second year teacher effectiveness. These findings may simply be unique to the select sample of corps members—work by Duckworth and colleagues indicates that corps members have significantly higher levels of grittiness than other young adults—and indicate that too much perseverance, perhaps being too perseverant to alter classroom practices in the face of challenges, may be detrimental to teacher effectiveness (Duckworth, Quinn, & Seligman, 2009). Given the hypothesized direction of these effects, however, future investigations with a more representative sample of teachers are necessary.

Regarding ratings of teacher competencies, for multiple evaluation standards higher measures of both cognitive (prior academic achievement) and non-cognitive (leadership and motivational ability) traits predicted significantly greater odds of rating above proficient. This indicates that non-cognitive characteristics likely contribute to the quality of teachers' practices. The traits predictive of high evaluation ratings, however, were not necessarily the same as those predictive of teacher effectiveness. Across outcomes results were congruent for prior academic achievement and motivational ability (positive effects for both outcomes), but differed for (1) organizational ability and respect, which did not significantly predict evaluation ratings and (2) leadership, which had negative value-added effects and positive evaluation results. This indicates that

principals' evaluation ratings may focus on different aspects of teaching quality than those identified by student achievement gains. Finally, several evaluation ratings predicted expected relationships between teacher traits and teacher competencies. For instance, teachers' (1) leadership ability predicted rating above proficient on standard one (teachers demonstrate leadership) and (2) prior academic achievement predicted rating above proficient on standard three (teachers know the content they teach). This suggests that specific teacher traits may manifest themselves in certain teacher behaviors—e.g. teachers with greater amounts of measured leadership ability take on greater leadership roles at a school, such as chairing a department or directing a student group—and that principals recognize such behaviors and rate teachers accordingly.

So how do research and policy move forward with these findings? Concerns regarding the representativeness of the sample—TFA is a highly select group working in low-income and low-performing schools—mean that this research is best thought of as a "proof of concept" to determine whether non-cognitive traits may matter and play a role in teacher quality policy. Given both the positive findings and the few unexpected, negative results, it would be useful for researchers to partner with teacher preparation programs and/or school districts to collect similar types of non-cognitive measures for a more generalizable sample of pre-service teachers. With such data researchers can ask: (1) what non-cognitive characteristics should be measured and tested (2) how reliably and cost-effectively are non-cognitive traits measured, especially for a larger sample of individuals (3) whether such measures predict future value-added effectiveness, the quality of teacher practice, and retention, both in the profession overall and in high-need schools (4) whether non-cognitive characteristics better explain differences in initial

teacher performance or how teachers develop on-the-job (5) whether the effects of specific non-cognitive traits are particular to certain grade levels or types of teachers (6) whether higher levels of non-cognitive characteristics may compensate for lower levels of measured cognitive ability (7) how these non-cognitive characteristics directly impact teacher behavior and (8) whether extremely high levels of a non-cognitive characteristic can have negative effects.

From a policy perspective, findings from these researcher-practitioner studies can facilitate results-aligned recruitment and selection/hiring practices by teacher preparation programs and school districts. For example, if measures of pre-service teachers' perseverance and grittiness predict student achievement gains, school districts could consider measures of these traits to hire, on average, more effective teachers. If multiple non-cognitive characteristics predict outcomes of interest, preparation programs and districts could create a composite index and select/hire individuals with scores above a cut-off. To be effective, however, this process requires sufficient numbers of applicants to preparation programs or districts to cull a select sample of those admitted or hired. Beyond selection, if non-cognitive traits are malleable, as research from other disciplines suggests, preparation programs and school districts can also structure training and professional development experiences to cultivate such characteristics. Overall, the potential exists to orient selection and development practices around non-cognitive characteristics directly linked to outcomes of interest.

This work is part of a promising, nascent research agenda to determine whether non-cognitive characteristics matter. In combination with other efforts, such as those to more deeply examine efficacious teacher preparation practices, the effects of school

environment on teacher development, or how specific teacher behaviors influence outcomes, research is moving closer to understanding what makes teachers effective.

REFERENCES

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago Public High Schools. *Journal of Labor Economics, 25*(1), 95-135.

Ames, C.A. (1990). Motivation: What teachers need to know. *Teachers College Record, 91*(3), 409-421.

Bacolod, M. (2007). Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality. *The Review of Economics and Statistics, 89*(4), 737-751.

Barrick, M., & Mount, M. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*(1), 1-26.

Borghans, L., Duckworth, A., Heckman, J., & Weel, B. (2008). The economics and psychology of personality traits. *IZA Discussion Paper No. 3333*.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & 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., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The effect of certification and preparation on teacher quality. *The Future of Children, 17*(1), 45-68.

Boyd, D., Grossman, P., Hammerness, K., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2012). Recruiting effective math teachers: Evidence from New York City. *American Educational Research Journal, 49*(6), 1008-1047.

Brophy, J. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology, 75*(5), 631-661.

Brophy, J. (1986). Teacher influences on student achievement. *American Psychologist, 41*(10), 1069-1077.

Chetty, R., Friedman, J., & Rockoff, J. (2011). The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood. *NBER Working Paper 17699.*

Clingman, J. & Fowler, R. (1976). The effects of primary reward on the I.Q. performance of grade-school children as a function of initial I.Q. level. *Journal of Applied Behavior Analysis, 9*(1), 19-23.

Clotfelter, C., Ladd, H., & Vigdor, J. (2005). Who teaches whom? Race and the distribution of novice teachers. *Economics of Education Review, 24*(4), 377-392.

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

Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682.

Clotfelter, C., Ladd, H., & Vigdor, J. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources, 45*(3), 655-681.

Cohen, D. (2011). *Teaching and its Predicaments.* Cambridge, MA: Harvard University Press.

Coleman, J., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F., et al. (1966). *Equality of Educational Opportunity.* Washington, DC: U.S. Government Printing Office.

Corcoran, S. (2007). Long-run trends in the quality of teachers: Evidence and implications for policy. *Education Finance and Policy, 2*(4), 395-407.

Danielson, C. (2007). *Enhancing Professional Practice: A Framework for Teaching.* Alexandria, VA: Association for Supervision and Curriculum Development.

Darling-Hammond, L., Holtzman, D., Gatlin, S.J., & Heilig, J.V. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach For America, and teacher effectiveness. *Education Policy Analysis Archives, 13*(42).

Decker, P., Mayer, D., & Glazerman, S. (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.

Dembo, M., & Gibson, S. (1985). Teachers' sense of efficacy: An important factor in school improvement. *The Elementary School Journal, 86*(2), 173-184.

Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417-440.

Dobbie, W. (2011). Teacher characteristics and student achievement: Evidence from Teach For America. Available from: http://www.people.fas.harvard.edu.

Duckworth, A., & Seligman, M. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*(12), 939-944.

Duckworth, A., Quinn, P., & Seligman, M. (2009). Positive predictors of teacher effectiveness. *The Journal of Positive Psychology, 4*(6), 540-547.

Duckworth, A., Grant, H., Loew, B., Oettingen, G., & Gollwitzer, P. (2011). Self-regulation strategies improve self-discipline in adolescents: Benefits of mental contrasting and implementation intentions. *Educational Psychology: An International Journal of Experimental Educational Psychology, 31*(1), 17-26.

Duckworth, A., Kirby, T., Tsukayama, E., Berstein, H., & Ericcson, K. (2011). Deliberate practice spells success: Why grittier competitors triumph in the National Spelling Bee. *Social Psychological and Personality Science, 2*(2), 174-181.

Farr, S. (2010). *Teaching as Leadership: The Highly Effective Teacher's Guide to Closing the Achievement Gap.* San Francisco: Jossey-Bass.

Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources, 47*(1), 765-794.

Goldhaber, D., & Liddle, S. (2011). The gateway to the profession: Assessing teacher preparation programs based on student achievement. *Center for Education Data and Research.*

Gordon, R., Kane, T., & Staiger, D. (2006). Identifying effective teachers using performance on the job. *The Brookings Institution: The Hamilton Project White Paper 2006-01.*

Greenwald, R., Hedges, L., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66*(3), 361-396.

Haberman, M. (1995). Selecting 'Star' teachers for children and youth in urban poverty. *Phi Delta Kappan, 76*(10), 777-781.

Hanushek, E. (2011). The economic value of higher teacher quality. *Economics of Education Review, 30*(3), 466-479.

Heckman, J.J. & Rubinstein, Y. (2001). The importance of non-cognitive skills: Lessons from the GED testing program. *American Economic Review, 91*(2), 145-149.

Heckman, J.J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900-1902.

Henry, G.T., Thompson, C.L., Brown, K., Cunningham, E., Kainz, K., Montrosse, B., Sgammato, A., & Yi, P. (2008). North Carolina high school resource allocation study. *Carolina Institute for Public Policy.* Available from: http://publicpolicy.unc.edu/research/publications-presentations-and-reports.

Henry, G.T., Bastian, K.C., & Smith, A.A (2012). Scholarships to recruit the 'best and brightest' into teaching: Who is recruited, where do they teach, how effective are they, and how long do they stay? *Educational Researcher, 41*(3), 83-92.

Henry, G.T., Bastian, K.C., Fortner, C.K., Kershaw, D.C., Purtell, K.M., Thompson, C.L., & Zulli, R.A. (2013). Teacher preparation policies and their effects on student achievement. In press, *Education Finance and Policy.*

Judge, T., Higgins, C. Thoreson, C., & Barrick, M. (1999). The Big Five personality traits, general mental ability, and career success across the life span. *Personnel Psychology, 52*(3), 621-652.

Kukla-Acevedo, S. (2009). Do teacher characteristics matter? New results on the effects of teacher preparation on student achievement. *Economics of Education Review, 28*(1), 49-57.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis, 24*(1), 37-62.

McKown, C., & Weinstein, R. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology, 46*(3), 235-261.

McREL. (2009). North Carolina Teacher Evaluation Process. Available from: http://www.ncpublicschools.org/docs/educatoreffect/ncees/instruments/teach-eval-manual.pdf.

Mischel, W., Shoda, Y., & Rodriguez, M. (1989). Delay of gratification in children. *Science, 244*(4907), 933-938.

Moffitt, T., Arseneault, L., Belsky, D., Dickson, N., Hancox, R., Harrington, H., Houts, R., Poulton, R., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *PNAS, 108*(7), 2693-2698.

Murray, H.B., & Staebler, B.K. (1974). Teacher's locus of control and student achievement gains. *Journal of School Psychology, 12*(4), 305-309.

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

Raymond, M.E., Fletcher, S., & Luque, J. (2001). Teach For America: An evaluation of teacher differences and student outcomes in Houston, Texas. Stanford, CA: CREDO, Stanford University. Available from http://credo.stanford.edu/downloads/tfa/pdf.

Ripley, A. (2010). What makes a great teacher? *The Atlantic.* Available from: www.theatlantic.com/magazine/archive/2010/01-what-makes-a-great-teacher/7841.

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

Rockoff, J., Jacob, B., Kane, T., & Staiger, D. (2011). Can you recognize an effective teacher when your recruit one? *Education Finance and Policy, 6*(1), 43-74.

Rockoff, J., & Speroni, C. (2011). Subjective and objective evaluations of teacher effectiveness: Evidence from New York City. *Labour Economics, 18*(5), 687-696.

Rosen, J.A., Glennie, E.J., Dalton, B.W., Lennon, J.M., & Bozick, R.N. (2010). *Noncognitive Skills in the Classroom: New Perspectives on Educational Research.* RTI Press Publication No. BI-0004-1009. Research Triangle Park, NC: RTI International. http://www.rti.org/rtipress.

Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinants of pupils' IQ gains. *Psychological Reports, 19*(1), 115-118.

Rothstein, J. (2010). Teacher quality in education production: Tracking, decay, and student achievement. *The Quarterly Journal of Economics, 125*(1), 175-214.

Staiger, D., & Rockoff, J. (2010). Searching for effective teachers with imperfect information. *Journal of Economic Perspectives, 24*(3), 97-117.

Stotko, E., Ingram, R., & Beaty-O-Ferrall, M.E. (2007). Promising strategies for attracting and retaining successful urban teachers. *Urban Education, 42*(1), 30-51.

Thompson, C.L., Brown, K.M., Townsend, L.W., Henry, G.T., & Fortner, C.K. (2011). Turning around North Carolina's lowest achieving schools (2006-2010). *Consortium for Educational Research and Evaluation—North Carolina.* Available from: http://publicpolicy.unc.edu/research/publications-presentations-and-reports.

Toch, T., & Rothman, R. (2008). Rush to judgment: Teacher evaluation in public education. *Education Sector Reports.* Available from: http://www.educationsector.org/sites/default/files/publications/RushToJudgment_ES_Jan08.pdf.

Wayne, A., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89-122.

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. *The New Teacher Project.* Available from: http://gcpstv.org/gcps-mainweb01.nsf/092DF14366B4598F8525788C0067CF48/$file/TNTPTheWidgetEffect.pdf.

Wiggins, G., & McTighe, J. (2005). *Understanding by Design.* Alexandria, VA: Association for Supervision and Curriculum Development.

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.

Table 1.1: Descriptive Information for the TFA Data

	Achievement	Critical Thinking	Leadership	Perseverance	Organizational Ability	Motivational Ability	Respect	Fit with TFA Mission
Mean and Std. Deviation	3.658 (0.301)	2.841 (0.386)	3.439 (0.967)	3.768 (0.549)	3.775 (0.708)	3.775 (0.535)	3.843 (0.723)	3.873 (0.595)
Achievement	1.000	---	---	---	---	---	---	---
Critical Thinking	0.099	1.000	---	---	---	---	---	---
Leadership	-0.160	-0.022	1.000	---	---	---	---	---
Perseverance	-0.136	-0.016	0.080	1.000	---	---	---	---
Organizational Ability	-0.060	-0.065	0.075	0.076	1.000	---	---	---
Motivational Ability	-0.057	-0.017	0.025	0.084	0.190	1.000	---	---
Respect	-0.054	-0.053	-0.037	0.074	0.091	0.095	1.000	---
Fit with TFA mission	-0.091	-0.062	-0.019	0.074	0.024	0.138	0.272	1.000
Praxis II Scores (std.)	0.271	0.186	-0.074	0.025	-0.127	-0.070	-0.013	-0.112

Note: The top portion of this table presents the means and standard deviations (prior to standardizing each trait) for each of the eight traits. The middle portion of this table displays the pair-wise correlations for the eight criteria measured by TFA during its selection process. Finally, the bottom portion of this table displays the correlations between teachers' Praxis II licensure test scores (a proxy for cognitive ability) and the eight skills and traits.

Table 1.2: The Selectivity of TFA Corps Members

Individual Teacher Characteristics		
Teacher Characteristics	**TFA Corps Members**	**All Other 1st and 2nd Year Teachers**
Age at entry into teaching	22.45 (1.71)	28.85 (9.03)
Percentage Female	74.93	77.88
Ethnicity Percentages		
White	81.08	79.78
Black	10.77	12.74
Hispanic	1.84	2.68
Other	6.31	4.79
Std. Praxis II Exam Scores	0.488 (0.657)	0.155 (0.709)
Barron's Ranking Percentages		
Not Competitive	2.02	1.92
Less Competitive	3.64	18.43
Competitive	16.46	55.73
Very Competitive	27.67	12.44
Highly Competitive	30.77	10.14
Most Competitive	19.43	1.34

Classroom and School Characteristics: Elementary Schools				
	Standardized Class Average	**Performance Composite**	**Classroom Percentage FRPL**	**School Percentage FRPL**
TFA Corps Members	-0.503 (0.303)	43.08 (10.63)	79.62 (16.93)	86.97 (14.02)
All Other 1st and 2nd Year Teachers	-0.197 (0.521)	57.22 (16.11)	62.74 (27.54)	61.52 (24.23)

Classroom and School Characteristics: High Schools				
	Standardized Class Average	**Performance Composite**	**Classroom Percentage FRPL**	**School Percentage FRPL**
TFA Corps Members	-0.548 (0.551)	63.72 (18.61)	62.98 (21.88)	63.34 (29.29)
All Other 1st and 2nd Year Teachers	-0.179 (0.636)	70.34 (15.78)	40.60 (27.10)	45.30 (21.74)

Note: The top half of this table displays individual characteristics for TFA corps members and all other 1st and 2nd year teachers (2007-08 through 2010-11). The bottom half of this table displays classroom and school characteristics (elementary and high schools) for tested-subject TFA corps members and all other 1st and 2nd year tested-subject teachers.

Table 1.3: Covariates Used in Value-Added Analyses

Student Covariates	Classroom and Teacher Covariates	School Covariates
1) Prior student test scores 2) Peer ability 3) Days absent 4) Structural mobility 5) Within year mobility 6) Between year mobility 7) Underage for grade 8) Overage for grade 9) Giftedness 10) Disability 11) Free or reduced-price lunch 12) Ethnicity 13) Gender 14) Currently limited English proficient 15) Was limited English proficient 16) Course indicators (HS only)	1) Class size 2) Heterogeneity of prior student performance 3) Advanced curriculum (secondary grades only) 4) Remedial curriculum (secondary grades only) 5) Out-of-field teaching 6) Teacher experience 7) Non-cognitive skills and traits	1) School size 2) School size squared 3) Total per-pupil expenditures 4) Average teacher supplement 5) Short-term suspension rate 6) Violent acts rate 7) Free and reduced-price lunch percentage 8) Race/ethnicity percentages

Note: All these covariates are included in value-added models. Single-year teacher experience indicators and all the school covariates are also included in the analyses of evaluation ratings.

Table 1.4: Do Non-Cognitive Skills & Traits Predict Teacher Effectiveness in Elementary Grades?

	Elementary Grades Mathematics			Elementary Grades Reading		
Skill or Trait	**1st Year Teachers**	**2nd Year Teachers**	**Combined**	**1st Year Teachers**	**2nd Year Teachers**	**Combined**
Prior academic achievement	-0.014 (0.025)	-0.039 (0.029)	-0.008 (0.020)	**0.033**$^{+}$ (0.018)	-0.033 (0.043)	0.017 (0.015)
Critical thinking	-0.003 (0.040)	0.068 (0.051)	0.008 (0.035)	-0.044 (0.028)	0.002 (0.031)	-0.032 (0.021)
Leadership	0.008 (0.027)	0.057 (0.038)	0.008 (0.025)	-0.022 (0.019)	0.036 (0.032)	0.005 (0.017)
Perseverance	-0.009 (0.031)	**-0.087**$^{+}$ (0.047)	-0.018 (0.024)	0.007 (0.020)	**-0.070**$^{+}$ (0.036)	0.002 (0.019)
Organizational ability	0.046 (0.037)	**0.071**$^{+}$ (0.036)	0.044 (0.029)	**0.058*** (0.023)	**0.070**$^{+}$ (0.035)	**0.041*** (0.018)
Motivational ability	0.016 (0.037)	**0.079*** (0.039)	0.026 (0.033)	-0.006 (0.020)	**0.079*** (0.036)	0.007 (0.016)
Respect for low-income students	0.014 (0.028)	0.021 (0.031)	0.016 (0.024)	0.008 (0.024)	0.046 (0.041)	0.027 (0.021)
Fit with TFA mission	-0.002 (0.016)	0.002 (0.034)	-0.002 (0.015)	0.012 (0.015)	-0.032 (0.038)	-0.007 (0.017)
Cases	3,072	1,599	4,671	4,024	2,301	6,325
TFA overall	**0.073*** (0.029)	**0.110**** (0.035)	**0.083**** (0.023)	0.023 (0.022)	0.002 (0.032)	0.009 (0.018)
Cases	105,736	123,603	229,339	141,847	167,636	309,483

Note: + Indicates significance at the p<0.10 level; * indicates significance at the p<0.05 level; ** indicates significance at the p<0.01 level.

Table 1.5: Do Non-Cognitive Skills & Traits Predict Teacher Effectiveness in High School?

	All High School End of Course Exams			STEM Subjects			Non-STEM Subjects		
Skill or Trait	**1st Year Teachers**	**2nd Year Teachers**	**Combined**	**1st Year Teachers**	**2nd Year Teachers**	**Combined**	**1st Year Teachers**	**2nd Year Teachers**	**Combined**
Prior academic achievement	-0.024 (0.020)	**0.050*** (0.025)	-0.002 (0.019)	-0.027 (0.021)	**0.073**** (0.027)	-0.000 (0.023)	-0.030 (0.023)	0.021 (0.044)	-0.035 (0.021)
Critical thinking	0.013 (0.029)	**-0.080**** (0.029)	-0.009 (0.025)	0.009 (0.035)	**-0.079*** (0.030)	-0.000 (0.034)	0.015 (0.035)	-0.052 (0.104)	-0.002 (0.031)
Leadership	-0.006 (0.026)	**-0.056**$^{+}$ (0.028)	-0.028 (0.024)	0.006 (0.036)	-0.032 (0.033)	-0.015 (0.029)	**-0.032**$^{+}$ (0.019)	**-0.116**** (0.042)	**-0.058**** (0.020)
Perseverance	-0.004 (0.021)	-0.033 (0.030)	0.002 (0.021)	-0.012 (0.026)	**-0.063*** (0.031)	-0.019 (0.027)	-0.002 (0.026)	0.033 (0.060)	**0.030**$^{+}$ (0.018)
Organizational ability	0.009 (0.023)	0.035 (0.024)	0.013 (0.018)	0.032 (0.030)	**0.085**** (0.032)	0.036 (0.026)	-0.036 (0.033)	-0.021 (0.057)	-0.021 (0.029)
Motivational ability	0.007 (0.020)	**0.053**** (0.020)	0.017 (0.016)	-0.007 (0.025)	0.025 (0.021)	0.002 (0.018)	-0.001 (0.032)	0.050 (0.056)	0.011 (0.035)
Respect for low-income students	**0.044**$^{+}$ (0.025)	**0.051*** (0.022)	**0.045*** (0.022)	0.020 (0.025)	**0.050**$^{+}$ (0.027)	0.024 (0.025)	0.038 (0.033)	**0.090**$^{+}$ (0.046)	**0.064**** (0.023)
Fit with TFA mission	-0.006 (0.025)	-0.028 (0.022)	-0.018 (0.022)	0.002 (0.039)	0.004 (0.029)	0.008 (0.031)	0.005 (0.019)	0.024 (0.038)	-0.021 (0.020)
Cases	10,409	7,145	17,554	7,287	5,029	12,316	3,122	2,116	5,238
TFA overall	**0.241**** (0.034)	**0.285**** (0.039)	**0.255**** (0.028)	**0.294**** (0.038)	**0.350**** (0.047)	**0.317**** (0.033)	**0.122**** (0.036)	**0.160**** (0.050)	**0.137**** (0.032)
Cases	169,071	230,180	339,251	99,130	130,615	229,745	69,941	99,565	169,506

Note: + Indicates significance at the p<0.10 level; * indicates significance at the p<0.05 level; ** indicates significance at the p<0.01 level.

Table 1.6: Do Non-Cognitive Skills & Traits Predict Teacher Evaluation Ratings?

Skill or Trait	Standard 1: Leadership	Standard 2: Respectful Environment	Standard 3: Content Knowledge	Standard 4: Facilitate Learning	Standard 5: Reflect on Practice
Prior academic achievement	1.139 (0.71)	**1.507*** (2.22)	**1.423**$^{+}$ (1.82)	**1.726**** (2.71)	1.309 (1.50)
Critical thinking	0.892 (-0.67)	1.225 (1.18)	1.090 (0.45)	0.971 (-0.18)	0.723 (-1.51)
Leadership	**1.454*** (2.10)	**1.429**$^{+}$ (1.79)	0.952 (-0.26)	**1.414**$^{+}$ (1.79)	1.314 (1.27)
Perseverance	1.169 (0.97)	1.260 (1.37)	1.340 (1.47)	1.042 (0.19)	0.739 (-1.47)
Organizational ability	1.049 (0.28)	1.061 (0.39)	1.012 (0.07)	1.111 (0.55)	0.916 (-0.47)
Motivational ability	1.180 (0.95)	**1.704*** (2.44)	**1.726*** (2.20)	1.286 (1.63)	**1.493*** (2.12)
Respect for low-income students	0.994 (-0.03)	0.803 (-1.43)	0.904 (-0.49)	0.869 (-0.92)	0.956 (-0.22)
Fit with TFA mission	1.128 (0.82)	**1.372*** (2.10)	0.855 (-0.75)	1.116 (0.70)	1.146 (0.79)
Cases	249	249	249	249	249

Note: Models include teacher experience controls and a rich set of school characteristics. The sample includes all corps members evaluated by their principal in 2010-11. Cells report odds ratios for being rated *above proficient* and z-scores. + indicates significance at the p<0.10 level; * indicates significance at the p<0.05 level; ** indicates significance at the p<0.01 level.

CHAPTER TWO

Teachers Without Borders: Consequences of Teacher Labor Force Mobility[1]

SUMMARY

In many states an initial response to teacher shortages was to grant reciprocal certification for individuals traditionally prepared out-of-state. To date, there has been little research investigating the effectiveness of these out-of-state prepared teachers; however, three hypotheses predict that teachers prepared out-of-state may be less effective: (1) labor markets force mobility on less qualified teachers; (2) out-of-state prepared teachers have less familiarity with the curricula, standards, and culture of the importing state; and (3) the attrition patterns of out-of-state prepared teachers. We examined the effectiveness of out-of-state prepared teachers using unique student level data from North Carolina and found that out-of-state prepared instructors are significantly less effective than in-state prepared and alternative entry peers in elementary grades mathematics and reading. After testing the three hypotheses above, evidence suggests that a lack of familiarity with the state's educational environment and attrition patterns help explain out-of-state prepared teachers' ineffectiveness.

[1] Kevin C. Bastian is the first author for this research and was responsible for the data management, analysis, organization and writing of this paper. Gary T. Henry is the second author and contributed to the framing, analysis plan, organization, and editing of the paper.

INTRODUCTION

In many states high population growth, teacher attrition, teacher retirements, and more employment opportunities for women have been responsible for teacher shortages (Bacolod, 2007; Common Core of Data; Ingersoll & Smith, 2003). These shortages, coupled with a need for more highly effective teachers to promote student achievement growth, have pushed many states to experiment with alternatives to solely licensing instructors prepared at in-state traditional education programs. Lateral/Alternative entry programs, which reduce barriers to employment by allowing individuals without teacher education credentials to complete requirements for certification while concurrently teaching, have been a common state policy response (Feistritzer, 2011; National Research Council, 2010; Shen, 1997). For example, from 2000-01 to 2009-10 the number of alternative entry teachers in North Carolina public schools increased 125 percent, from 6,626 to 15,028, and the percentage of alternative entry teachers in the state's workforce rose from 7.79 to 14.87 (authors' analysis).

Through licensing agreements with national teacher accreditation and certification associations, such as the National Council for the Accreditation of Teacher Education or the National Association of State Directors of Teacher Education and Certification, a frequently used but little studied alternative approach has been for states to grant more reciprocal teacher certification licenses and expand the number of traditionally prepared teachers from other states. This policy broadens the pool of potential teachers by facilitating the interstate movement of experienced teachers and teacher candidates, especially from states that over-produced instructors in their education programs to those states in need of additional teachers. In North Carolina, for example, the states that have

contributed the largest share of the out-of-state prepared teacher pool are New York, Pennsylvania, Ohio, and Michigan—all states with recent decreases in their student populations (Common Core of Data)—and over the past decade, the number of out-of-state prepared instructors has increased 36 percent, from 21,316 to 29,006 (authors' analysis).

Despite these three options (in-state, out-of-state, and alternative entry) for staffing schools, prior research on teacher effectiveness has generally combined in-state and out-of-state traditionally prepared teachers into a single category and compared the effectiveness of traditionally prepared teachers with that of alternatively prepared instructors (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; Clotfelter, Ladd, & Vigdor, 2007, 2010; Kane, Rockoff, & Staiger, 2008). From a policy perspective, classifying in-state and out-of-state traditionally prepared teachers together ignores differential financial costs associated with these groups of instructors, including those of preparation, recruitment, development, and replacement which make distinctions between the categories important (Alliance for Excellent Education, 2004). Furthermore, combining those traditionally prepared in-state with out-of-state prepared instructors ignores potential differences in teacher preparation and labor markets across states which may affect teachers' effectiveness (National Research Council, 2010).

Both theoretical and prior research evidence suggests three hypotheses predicting that out-of-state prepared teachers may be less effective than their in-state prepared peers: (1) teacher candidates with lower levels of human capital may need to be more mobile to find employment (Boyd, Lankford, Loeb, & Wyckoff, 2005; Reininger, 2012); (2) differences in state curricula, standards, and culture may make out-of-state prepared

teachers less familiar with the educational environment of the importing state, and therefore, less effective in raising student achievement (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009); and (3) out-of-state prepared teachers who acquire human capital through on-the-job experience may become more competitive for positions back in their state of origin, causing high rates of teacher turnover and the potential for differential attrition (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006).

In this paper we examine the effectiveness of early-career out-of-state prepared teachers and address each hypothesis by answering the following research questions:

1) How does the effectiveness of traditionally prepared out-of-state teachers compare to that of individuals traditionally prepared in-state, or admitted into the profession through alternative entry programs?

2) If effectiveness differences exist between out-of-state prepared teachers and the groups specified above, what accounts for those performance disparities—lower levels of human capital, less familiarity with the educational environment, and/or the attrition patterns of out-of-state prepared instructors?

By way of preview, we find that out-of-state prepared teachers consistently underperform in-state traditionally prepared elementary school teachers in both mathematics and reading. In addition, out-of-state prepared teachers are significantly less effective than alternative entry instructors across all model specifications. Upon testing hypotheses to account for these findings, our results indicate that out-of-state prepared teachers' (1) lack of familiarity with the state's education environment and (2) high rates of turnover, coupled with the ineffectiveness of departing out-of-state prepared teachers, help explain their poor performance.

In the following sections we summarize the research investigating the effects of teacher preparation and detail our research-based hypotheses for why out-of-state prepared teachers may be less effective. We then discuss the data and methods used for this research. Next, we present the results for both research questions, and finally, we conclude with a discussion of potential policy responses.

PRIOR RESEARCH ON THE EFFECTS OF TEACHER PREPARATION

Recent changes in teacher labor markets—due to increased demand for teachers, particularly in certain high-need subject areas and schools—and concerns regarding low levels of student performance have pushed policymakers to open more alternative routes into the teaching profession (Shen, 1997). For instance, in 1998-99 the number of new teachers entering the profession through alternative pathways stood at 10,000. By 2005-06, that number had increased five-fold, and in 2009-10 approximately forty percent of teachers entering the profession within the last five years had done so through alternative routes (Feistritzer, 2011). This rapid increase in the alternatively prepared teaching population has provided researchers the opportunity to examine the efficacy of traditional teacher preparation, relative to alternative preparation, and within the past decade a number of studies have compared the effectiveness of traditionally and alternatively prepared/certified teachers.

Overall, this body of research has generated two broad findings. First, teachers holding regular certification/traditional preparation appear to be more effective in the early stages of their careers (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; Clotfelter, Ladd, & Vigdor, 2007, 2010; Goldhaber & Brewer, 2000; Henry, Bastian, Fortner, Kershaw, Purtell, Thompson, & Zulli, 2013; Kane, Rockoff, & Staiger, 2008).

These returns to certification/preparation fade quickly, however, such that the efficacy of the credential as a signal of teacher quality is limited (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006). For example, high school teachers holding an initial or continuing license in North Carolina public schools are more effective than those currently holding lateral entry licenses, but compared with those teachers who previously held a lateral entry license, no differences in effectiveness exist (Clotfelter, Ladd, & Vigdor, 2010). Second, there is more variation in teacher effectiveness within preparation categories than between them, meaning factors outside preparation may better determine a teacher's classroom success (Boyd, Goldhaber, Lankford, & Wyckoff, 2007; Henry et al., 2013; Kane, Rockoff, & Staiger, 2008). Despite the tenor of these findings four methodological or sample issues provide justification for further research on the relative effectiveness of teachers prepared through different programs or entering through different teacher preparation portals.

First, most of the prior research is based on teachers' certification status at a particular point in time, not on their preparation prior to beginning teaching. Preparation is the fixed education and training an individual brings with them into the teaching profession, while certification varies over time depending upon the grade, course, the types of students being taught, and teachers' professional development experiences. Two teachers, one who entered the profession with regular certification through a traditional preparation program, and another who entered with alternative certification and acquired regular certification through additional coursework and passing required tests, can hold the same certification status after a few years of experience. Therefore, research that draws conclusions about the value of traditional teacher preparation, while using

certification status as a proxy for traditional preparation, may not be accurately estimating preparation effects (Goldhaber & Brewer, 2000).

Second, research documents significant amounts of heterogeneity in preparation components and requirements within traditional and alternative portals, yet many studies only include these two broad categories in their analyses (Darling-Hammond, Berry, & Thoreson, 2001; Greenberg, Pomerance, & Walsh, 2011; National Research Council, 2010). Third, to best measure the influence of teacher preparation, research should focus on teachers early in their careers, when the effects of preparation are the strongest (Goldhaber & Liddle, 2011, Henry et al. 2013). To date, only a few studies limit their analyses to this early-career sub-sample (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006; Henry et al., 2013), while others draw inferences about the value of teacher preparation based upon the effectiveness of teachers more than a decade removed from formal training (Clotfelter, Ladd, & Vigdor, 2007, 2010; Goldhaber & Brewer, 2000).

Finally, due to teacher selection into training portals (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2006) and the non-random attrition of teachers from the profession (Chingos & Peterson, 2011; Goldhaber, Gross, & Player, 2011; Henry, Bastian, & Fortner, 2011; Krieg, 2006), great care must be taken when interpreting preparation estimates. Without a measure of teachers' academic ability or general human capital prior to entering a teacher preparation program, estimates of teachers' effectiveness combine the effects of selection and preparation; likewise, if teachers exit the profession from certain portals at greater rates than others, or if more or less effective teachers exit at greater rates from certain portals, then these estimates combine the effects of teacher

preparation and longevity (Henry, Bastian, & Fortner, 2011; Henry, Fortner, & Bastian, 2012).

While there is a burgeoning research literature on performance disparities between regular and alternatively certified teachers, and some research contrasting traditionally and alternatively prepared teachers, there is currently a dearth of literature focusing on out-of-state prepared instructors. To date, two studies have separately examined the effectiveness of out-of-state prepared teachers: (1) Henry and colleagues, 2013, in North Carolina, and (2) Goldhaber and Liddle, 2011, in Washington State. Henry et al. found that teachers prepared out-of-state were significantly less effective than traditionally prepared in-state teachers in elementary grades mathematics and reading, where out-of-state prepared teachers are highly concentrated, and in high school mathematics and science. While this work represents an important advance in teacher preparation research, this study did not: (1) compare the effectiveness of out-of-state prepared teachers with that of teachers from any other route into the profession other than in-state prepared teachers or (2) investigate explanations for these performance disparities. These points will be addressed in this study. In contrast, Goldhaber and Liddle return few effectiveness differences between teachers in Washington prepared at in-state education programs and out-of-state prepared teachers.

In the next section we lay out our hypotheses concerning disparities in performance for out-of-state prepared teachers that motivated this study.

EXPLANATIONS OF PERFORMANCE DISPARITIES

The Quality of Imported Teachers

Prior research indicates that teachers have geographically small labor markets, preferring to work close to their hometown and/or undergraduate institution (Boyd, Lankford, Loeb, & Wyckoff, 2005; Reininger, 2012). Additionally, research shows that teachers with higher levels of human capital are, on average, more effective (Clotfelter, Ladd, & Vigdor, 2007, 2010; Dobbie, 2011; Goldhaber, 2007; Greenwald, Hedges, & Laine, 1996; Henry, Bastian, & Smith, 2012; Rockoff, Jacob, Kane, & Staiger, 2011). Taken together, we hypothesize that in states/labor markets where the supply of teachers exceeds demand—especially in low-growth states where education programs prepare more teachers than can be absorbed into the workforce—if the labor market operates with reasonable efficiency the teachers with higher levels of human capital will be hired locally and the teachers with lower levels of human capital will be forced to broaden their job search to states experiencing teacher shortages. Quite simply, labor market forces, such as insufficient levels of local demand for new teachers and teacher preferences for proximity to home, may push less-skilled and lower human capital teachers to seek teaching positions in states, such as North Carolina, with teaching shortages.

To empirically examine this hypothesis we test whether a measure of teacher human capital mediates the effectiveness differences between out-of-state prepared and in-state prepared and alternative entry teachers. From the view of the importing state, the relevant question is the level of human capital of the imported out-of-state prepared teachers compared with the human capital of the in-state prepared and alternative teachers hired in the importing state. To select a mediator for this analysis, a measure of

individual human capital taken prior to entry into the teaching profession that is significantly associated with student achievement gains is required. From these criteria we selected a standardized, composite measure of all available teacher test scores—SAT/ACT, Praxis I exams, Praxis II exams—which is shown in prior work to be significantly associated with teachers' ability to increase students' achievement (Clotfelter, Ladd, & Vigdor, 2007, 2010; Goldhaber, 2007).[2] If out-of-state prepared teachers are no longer significantly less effective or the magnitude of the effectiveness differences is markedly reduced after the inclusion of this human capital indicator, this would suggest that the quality of imported teachers helps explain the underperformance of out-of-state prepared instructors.

Teachers' Lack of Familiarity with the Importing State

As a result of the state standards/accountability movement and No Child Left Behind, each state has unique standards, curriculum, and assessments on which their educational systems are built.[3] Importantly, the differences between states in these standards and assessments are substantial, and to the extent that teacher preparation programs are regulated by states, in-state colleges and schools of education may structure their course content and academic requirements to enable their graduates to become more familiar with the state curricula and academic content than individuals prepared elsewhere (Carnoy & Loeb, 2002; Porter, Polikoff, & Smithson, 2009). Through student

[2] In prior research another indicator of teacher human capital, the Barron's ranking for a teacher's undergraduate institution, has sometimes been positively associated with student achievement gains (Clotfelter, Ladd, & Vigdor, 2007, 2010). However, in our preliminary analyses the Barron's ranking was not correlated with achievement gains after adjusting for other covariates. Thus, we rely only on teacher test scores to test our first hypothesis.

[3] Through the recent adoption of the Common Core, state standards and assessment systems will soon become more similar. During the study period for this analysis (2005-06 through 2009-10), however, there were significant differences between states.

teaching and other field experiences, pre-service teachers at in-state colleges and schools of education will also enjoy more opportunities to engage in teaching practice in educational environments—types of students, curriculum/content, and schools—similar to their future, in-service classroom placements. Based on recent research which suggests that early-career instructors benefit from greater pre-service exposure to the school environments in which they will teach and the academic content they are expected to teach, we hypothesize that out-of-state prepared teachers will be less familiar with North Carolina's educational environment and therefore, less effective than in-state prepared teachers (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009). Out-of-state prepared teachers may also underperform in-state prepared and alternative entry instructors, many of whom have been long-term North Carolina residents, due to a lack of familiarity with the importing state's culture and students.

We test this lack of familiarity hypothesis in two ways. First, based on the theory that teachers in adjacent states might be more familiar with the importing state given regional similarities, we separate out-of-state prepared instructors into two groups—those entering North Carolina with preparation from a university in a contiguous state and those entering North Carolina with preparation from a university in a non-contiguous state.[4] For this hypothesis to hold, we expect: (1) teachers prepared in non-contiguous states to be less effective, particularly underperforming their in-state prepared peers who have greater exposure to the state's schools, curricula, and culture; and (2) teachers prepared in contiguous states to perform similarly to their in-state prepared and alternative entry peers. Second, we compare the effectiveness of out-of-state prepared

[4] This test of the lack of familiarity hypothesis is particularly salient since North Carolina imports over 50% of its early-career out-of-state prepared teachers from four states in different parts of the country—New York, Pennsylvania, Michigan, and Ohio.

teachers in their first year teaching (when they should possess less knowledge of the state's educational environment) with that of first year in-state prepared and alternative entry teachers. For this hypothesis to hold, we expect first year out-of-state prepared teachers to be less effective, particularly underperforming their in-state prepared peers whose traditional training increased their familiarity with the state's educational environment.

Teacher Turnover and Differential Attrition

If, as the first hypothesis and the research cited there suggests, teachers have strong preferences to teach in their home state but are forced through competition to seek teaching positions in other states, we hypothesize that as out-of-state prepared teachers acquire more human capital through experience in the classroom, they will become more attractive for open teaching positions in their states of origin. If this is true, out-of-state prepared teachers can be expected to leave North Carolina public schools at significantly greater rates than teachers with other forms of preparation. Since teachers have been shown to become more effective during their first few years on the job, this turnover may adversely influence student achievement, reduce school stability, substantially increase teacher recruitment/replacement costs, and contribute to large numbers of novice teachers in the workforce (Alliance for Excellent Education, 2004; Henry, Bastian, & Fortner, 2011; Henry, Fortner, & Bastian, 2012; Ronfeldt, Loeb, & Wyckoff, 2013). To establish whether turnover is higher among out-of-state prepared teachers than in-state prepared or alternative entry instructors, we compare teachers' odds of leaving North Carolina public schools.

In addition to the potential adverse effects of overall teacher turnover rates, differential attrition may also contribute to out-of-state prepared teachers' ineffectiveness. If the most effective out-of-state prepared teachers exit the state's public schools, the quality of the remaining out-of-state prepared population would be reduced. Conversely, if the least effective out-of-state prepared teachers exit North Carolina, then the poor performance of out-of-state prepared teachers may actually represent an upwardly biased estimate of their effectiveness. We test this differential attrition hypothesis by examining the effectiveness of out-of-state prepared teachers who: (1) will leave North Carolina public schools before beginning their sixth year of teaching; and (2) will not return to North Carolina public schools in the following school year (last year). Finally, we examine the extent to which the effort and effectiveness of out-of-state prepared teachers decreases in their last year of employment in North Carolina public schools, a phenomenon known as the Ashenfelter Dip (Ashenfelter, 1978).

DATA AND SAMPLE

The main objectives for this study were to estimate the comparative effectiveness of out-of-state prepared teachers in elementary school mathematics and reading and to investigate three potential explanations for out-of-state prepared teachers' underperformance. This required developing and applying a teacher preparation coding scheme, matching students to their classroom teachers, building longitudinal analysis files with student, teacher, classroom, and school characteristics, and estimating effects. The following sections detail our classification of teachers, study sample, and covariates used in the analyses.

Classification of Teachers

Our classification of teachers is based on a combination of the formal education and specific preparation teachers possessed when they began teaching. This means we grouped teachers into exclusive and fixed categories according to their formal preparation—earning a degree or completion of a certificate program—most proximate to entering the profession, whether they completed their preparation in-state or in another state, and whether they had completed all requirements for initial certification. In total, we created three policy relevant teacher preparation categories for this analysis: (1) out-of-state prepared (fully certified upon first entering the teaching profession after earning an undergraduate/graduate degree or completing a licensure/certification program); (2) in-state prepared (fully certified upon first entering the teaching profession after earning an undergraduate/graduate degree or completing a licensure/certification program); and (3) alternative entry (not fully certified upon first entering the teaching profession.[5]

In order to classify North Carolina elementary school teachers into these groups, we relied on administrative data from three sources: (1) institutional data from the University of North Carolina General Administration (UNCGA) that identified in-state publicly prepared teachers at the undergraduate, graduate, and licensure/certificate level; (2) teacher education, licensure audit, and certified salary files from the North Carolina Department of Public Instruction (NCDPI); and (3) identifiers of Teach For America corps members in the state. From these datasets we employed several key pieces of information to classify teachers. First, we calculated the year an individual began

[5] Approximately 2% of the sample was unclassifiable because: (1) they did not have a college graduation year in the administrative data; (2) their highest degree earned prior to beginning teaching was less than a bachelor's degree; or (3) administrative data recorded the person teaching more than one year prior to their earliest graduation year. We retained these unclassifiable teachers in analyses but do not report their effects. Results for unclassifiable teachers are available upon request.

teaching, which facilitated the identification of an individual's most proximate preparation prior to entry. Second, using the NCDPI licensure audit file we determined the basis for a teacher's original teaching license. If this initial basis indicated that a teacher had not completed all licensure requirements upon first entering the profession, we classified teachers as alternative entry (without regard to the state from which they received their final degree before beginning to teach). Finally, using the UNCGA and the NCDPI education files, we determined an individual's graduation year, degree level, and degree origin and assigned individuals to a single teacher preparation category. For example, if an individual earned an undergraduate degree from an out-of-state institution, did not earn any additional degrees (at an in-state institution) before entering the profession, and had completed all licensure requirements prior to entry, we classified them into the out-of-state prepared category. If an individual earned multiple degrees prior to beginning teaching, we classified them according to the degree most proximate to their entry into the profession. We believe this coding scheme and its focus on the fixed preparation and training of teachers as they enter the profession represents an advance over coding schemes based on teachers' certification status and allows us to accurately estimate the effects of preparation.[6]

Study Sample

The data for this analysis span the 2005-06 through the 2009-10 school years and are limited to teachers with less than five years teaching experience. We restrict our sample to these early-career teachers for three reasons. First, we believe teachers are most likely to display measurable and relevant preparation influences early in their careers,

[6] Here, estimates of preparation would include both the training and formal education received by participants and the selection of participants into particular preparation pathways or portals.

before on-the-job learning overwhelms effects. Second, for states considering policy mechanisms to improve the quality of their teacher workforce, individuals who recently entered the profession entail particular significance since they comprise more than twenty-five percent of the teaching population (Ingersoll & Merrill, 2010). Finally, we limit our sample because out-of-state prepared teachers are highly concentrated in elementary schools, representing nearly 37 percent of the early-career, tested-grades (3-5) instructors (See Table 2.3 for unique teacher counts from models).

For this analysis, the key data feature is our use of actual classroom rosters, which allowed us to validly match students to approximately ninety-three percent of individual instructors over the five-year study period, construct classroom level covariates, and account for multiple teachers within a subject-year for a given student. Numerous other student, teacher, classroom, and school characteristics were merged into these files and used in various model specifications to account for factors influencing student achievement outside the control of a teacher preparation category. In total, across elementary school mathematics and reading models, we analyzed 886,865 test scores, 447,347 unique students, and 12,192 unique teachers.

Dependent Variables and Covariates

For this analysis students' prior and current test score performance is based on the North Carolina grade three pre-test and the End-of-Grade (EOG) mathematics and reading exams in grades 3-5.[7] These EOG exams are criterion-referenced—based on the North Carolina Standard Course of Study objectives for each course—vertically equated across years to allow for meaningful comparisons, and have been rigorously analyzed to

[7] Before the 2009-10 school-year North Carolina discontinued the 3rd grade pretest. Therefore, value-added outcomes are not available for 3rd grade teachers in that year.

ensure valid psychometric properties (North Carolina Department of Public Instruction, 2011). To remove any secular trends, all tests were standardized within subject, grade, and year, such that a standardized value of zero represents the average score for that subject-year, and consecutive standardized values of zero for a particular student indicate that she made gains equivalent to the average student. Additionally, we included year fixed effects in our model specifications. The focal variables used in this analysis are indicator variables for in-state prepared and alternative entry teachers in comparison with out-of-state prepared teachers (reference category). To adjust for factors influencing teacher effectiveness outside the control of teacher preparation, including potential distributional differences by teacher preparation type, models also include a rich set of student, classroom, teacher, and school covariates, as listed in Table 2.1. Finally, depending on the model specification or hypothesis test, some models drop or include control variables for particular analyses.

ANALYSIS PLAN

Teacher Value-Added

The goal of this analysis is to estimate the relative effectiveness of out-of-state prepared teachers compared to in-state prepared and alternative entry instructors. To do this we utilize our extensive administrative database and specify an ordinary least squares (OLS) value-added model with a rich set of student, classroom, teacher, and school covariates (Table 2.1). We use cluster-adjusted standard errors at the school-year level to account for the nesting of students and teachers within schools that, if left unadjusted, could result in reduced standard errors and significance tests that produce false positives. The equation used to estimate teacher preparation category effects is as follows:

$$Y_{ijst} = \beta_0 + \beta_1 Y_{it-n} + \beta_2 Instate + \beta_3 Alternative + \beta_x X_{ijst} + \beta_z Z_{jst} + \beta_w W_{st} + \varepsilon_{ijst} \quad (1)$$

where Y_{ijst} is the test score for student *i* in classroom j in school *s* at time *t*;

β_2 and β_3 estimate the average effect of in-state prepared and alternative entry instructors relative to out-of-state prepared teachers;

In-state and *Alternative* are indicator variables that equal 1 if the teacher entered teaching through that category and 0 if not;

Y_{it-n} represents the prior test scores for student *i*;

X_{ijst} represents a set of individual student covariates;

Z_{jst} represents a set of classroom and teacher covariates;

W_{st} represents a set of school covariates;

and ε_{ijst} is a disturbance term representing all unexplained variation.

In response to our preferred rich covariate adjustment model, the fundamental question is how well it controls for endogeneity threats. Recent studies assessing alternative identification strategies have shown that covariate adjusted estimates, when rich covariates are available, substantially reduce bias in effect estimates in comparison to estimates from a randomized control trial (Bifulco, 2012; Glazerman, Levy, & Myers, 2003; Shadish, Clark, & Steiner, 2010). Furthermore, we prefer the OLS approach because rich covariate adjustment models estimate teacher preparation category effects based on the entire (statewide) sample of teachers, rather than the more limited sample of within-unit variation in fixed effects approaches. However, if the non-random assignment of teachers to students or unmeasured school factors, such as school leadership quality, affect student test performance and are (1) correlated with the preparation categories and (2) omitted from models, then fixed effects will produce preferred, internally valid estimates of teacher preparation effectiveness.

Therefore, to assess the robustness of our preferred OLS value-added model, we employ three fixed effects specifications—school fixed effects, school-by-year fixed effects, and student (levels model) fixed effects. Here, the school fixed effects limit comparisons to students and teachers within the same school, thereby eliminating any uncontrolled, time-invariant school factors that may influence estimates, while the school-by-year fixed effects restrict teacher preparation comparisons to students and teachers in the same school and year to control for unmeasured school and temporal trends. Our student fixed effects model is a levels (no prior test score) specification that uses students as their own control and compares students' test score outcomes (deviation from the students' mean scores standardized by grade and year) when taught by an out-of-state prepared teacher to outcomes when instructed by an in-state prepared or alternative entry teacher. These fixed effects models continue to include a rich set of student, classroom, teacher, and school covariates to isolate the effect of preparation categories—we exclude time-invariant student characteristics from the student fixed effects models and school covariates from the school-by-year fixed effects models. Because these fixed effects approaches only identify coefficients based on within-unit (school, school-by-year, or student) variation, the results table (Table 2.3) for our first research question includes counts of unique teachers contributing to the preparation category estimates. Furthermore, all results tables provide observation counts for the student test records that contributed to the teacher preparation estimates. For instance, observation counts for school fixed effects models exclude schools in which no out-of-state prepared teachers worked or only out-of-state prepared teachers worked during our study period. Due to the small sample of student test records identifying estimates in the

student fixed effects models—see Table 2.3 for counts—we only use the OLS, school fixed effects, and school-by-year fixed effects models to test our three research-based hypotheses.

Teacher Turnover

To test part of our third hypothesis, we examine teacher turnover using a logistic regression framework with last year (not returning to North Carolina public schools in the following school year) as the dependent variable, out-of-state prepared teachers as the reference category, and a set of classroom, teacher, and school covariates to control for differences in employment context that may influence teacher persistence in North Carolina public schools. The equation for this specification is as follows:[8]

$$\Pr(lastyear_{it} = 1 = \frac{\exp(teacherprep_i\alpha + Class_{it}\beta + School_{it}\tau)}{1 + \exp(teacherprep_i\alpha + Class_{it}\beta + School_{it}\tau)} \quad (2)$$

where $lastyear_{it}$ is a binary indicator for whether a teacher returns to North Carolina public schools in the following school year;

$teacherprep_i$ is a set of teacher preparation indicators for in-state and alternative entry instructors in reference to out-of-state prepared teachers;

and $class_{it}$ and $school_{it}$ are a set of classroom/teacher and school contextual factors that may influence teacher persistence.

FINDINGS

Descriptive Information

Before reviewing the teacher preparation results from rich covariate and fixed effects models, we briefly present descriptive information regarding students' academic

[8] In addition to this logistic regression, we also performed logit regressions with school and school-by-year fixed effects. Results from these models were very similar to our presented findings in Table 2.7.

achievement and economic status in the schools and classrooms in which our sample of teachers work. Here, we use standard independent sample t-tests to determine whether there are statistically significant differences in the school/classroom environments of out-of-state prepared teachers and in-state prepared or alternative entry instructors. Examining Table 2.2, the data show that in comparison to alternative entry teachers, early-career out-of-state prepared instructors teach: (1) students with higher average prior test scores; (2) in classrooms where fewer students qualify for subsidized lunches; and (3) in schools where more students pass their EOG tests and fewer students qualify for subsidized lunches. In comparison to in-state prepared teachers, early-career out-of-state prepared instructors also teach in classrooms and schools with a lower percentage of students qualifying for subsidized lunches. Overall, the data suggest that, relative to in-state prepared or alternative entry instructors, out-of-state prepared teachers are not distributed into more challenging working conditions (in fact the reverse is more consistent with the data), as measured by student performance and economic disadvantage indicators, which might adversely influence their effectiveness.

How Effective are Out-of-State Prepared Teachers?

For our preferred rich-covariate adjustment model, Table 2.3 shows that both in-state prepared and alternative entry teachers significantly outperform out-of-state prepared teachers in mathematics and reading. In-state prepared instructors are 0.024 and 0.012 standard deviations more effective in mathematics and reading, respectively; alternative entry teachers are 0.030 and 0.013 standard deviations more effective. Using school, school-by-year, or student fixed effects to adjust for the sorting of students and teachers or other unmeasured school characteristics does not alter the substance of these

results. Out-of-state prepared teachers significantly underperform their in-state prepared and alternative entry peers across all model specifications.

While the significant results in Table 2.3 are not large, approximately one to three percent of a standard deviation, depending upon the subject, this effect size is comparable to findings from other early-career teacher preparation research (Boyd et al., 2006; Kane, Rockoff, & Staiger, 2008). To make these estimates more tangible, we used the average gains in EOG scale score points between elementary grades and the average standard deviation on elementary grades EOG tests to convert the effects in Table 2.3 into equivalent days of student learning in a 180 day school calendar. Here, in comparison to students instructed by out-of-state prepared teachers, the in-state prepared effects of 0.024 and 0.012 in the rich covariate adjustment mathematics and reading models are worth approximately 6 and 3.5 additional days of student learning, respectively; the alternative entry effects of 0.030 and 0.013 are worth approximately 7.5 and 3.7 additional days of student learning in mathematics and reading, respectively (Henry, Thompson, Fortner, Bastian, & Marcus, 2011). In practical terms for statewide student achievement, the magnitude of these effects must be considered alongside the size of the teacher preparation category. As shown in the bottom panel of Table 2.3, out-of-state prepared teachers are the second largest source of elementary school teachers in North Carolina, comprising approximately 37 percent of the early-career tested-grades teacher workforce. During this five year study period these teachers taught nearly 200,000 students in grades 3-5, meaning, for example, that replacing all out-of-state prepared teachers with in-state prepared teachers would be equivalent to 1.2 million days of additional student learning in elementary grades mathematics. Consequently, the poor

performance of out-of-state prepared teachers has widespread effects. In the following sections we attempt to explain the performance of out-of-state prepared teachers.

Why Do Out-of-State Prepared Teachers Underperform?

The Quality of Imported Teachers: To empirically examine our first hypothesis concerning the human capital of imported instructors, we test whether an indicator of teacher quality—a standardized, composite measure of all available teacher test scores[9]—mediates the effectiveness differences between out-of-state prepared and in-state prepared and alternative entry teachers shown in Table 2.3. Following the mediation procedures set forth in Shrout and Bolger, we began by determining whether our teacher test score measure was significantly associated with student achievement (Shrout & Bolger, 2002). Results presented in the top panel of Table 2.4—from a model with a rich-set of student, classroom/teacher, and school covariates, but without teacher preparation covariates—show that a composite measure of teacher test scores strongly predicts student achievement gains (In the elementary grades reading school-by-year fixed effects model the p-value was 0.059). This finding corroborates prior research and indicates that, on average, teachers with higher levels of human capital are more effective (Clotfelter, Ladd, & Vigdor, 2007, 2010; Goldhaber, 2007). Next, to determine whether this human capital indicator mediates the effectiveness differences between our teacher preparation categories, we ran our rich covariate and fixed effects models (school and school-by-year) and included the teacher test score variable. The bottom panel of Table

[9] Approximately 18% of the teachers in our study sample were originally missing a value for the composite teacher test score variable. In order to run our mediation analyses on the same sample of teachers and students we created a dataset with all available information about our sample teachers and the classrooms and schools in which they work and used the SAS proc mi (multiple imputation) command to impute the missing teacher test score values. Findings from a non-imputed mediation analysis are comparable to the presented results and available upon request.

2.4 shows that even after controlling for teacher test scores, in-state prepared and alternative entry teachers remain significantly more effective than out-of-state prepared teachers. Comparing the teacher preparation category results in Table 2.4 with those in Table 2.3, we find no evidence of mediation. All of the in-state prepared coefficients are equivalent between the tables, while the alternative entry effects in math are slightly larger with the mediator included. With these findings we reject our first hypothesis as an explanation for out-of-state prepared teachers' underperformance. Human capital differences between out-of-state prepared teachers and those in-state prepared and alternative entry instructors working in North Carolina do not appear to explain why out-of-state prepared teachers underperform. It is still possible, however, that human capital differences do exist between out-of-state prepared teachers working in North Carolina and their peers who were hired in their states of origin.

Teachers' Lack of Familiarity with the Importing State: We test our lack of familiarity hypothesis in two ways. First, we separated out-of-state prepared teachers into two groups—those entering North Carolina with preparation from a contiguous state (Virginia, Tennessee, Georgia, and South Carolina) and those entering with preparation from a non-contiguous state—and ran separate models comparing contiguous and non-contiguous out-of-state prepared instructors (as the reference group) to our other teacher preparation categories. Here, we hypothesized that teachers prepared in contiguous states, due to a greater familiarity with North Carolina's educational environment and culture, would perform comparably to in-state prepared and alternative entry instructors, while teachers prepared in non-contiguous states would continue to underperform. As shown in Table 2.5, results for in-state prepared teachers in the elementary school reading models

substantiate our hypothesis. In the rich covariate, school, and school-by-year fixed effects specifications, out-of-state prepared teachers from contiguous states perform comparably to in-state prepared teachers (top panel of Table 2.5), while out-of-state prepared teachers from non-contiguous states continue to be less effective than in-state prepared teachers (bottom panel of Table 2.5). While our hypothesis holds for alternative entry instructors in the rich covariate reading models, alternative entry teachers remain significantly more effective than both contiguous and non-contiguous out-of-state prepared teachers in the fixed effects specifications. Results from the elementary school mathematics models do not support our hypothesis—whether prepared in states contiguous or non-contiguous to North Carolina, out-of-state prepared teachers remain less effective than in-state prepared and alternative entry instructors. Because most elementary school classrooms in North Carolina are self-contained, meaning the same teacher is responsible for both mathematics and reading instruction, these in-state prepared reading findings suggest that the importance of regional familiarity differs across the two subjects. Whether these reading results are due to differences in the curriculum across states or unfamiliarity with the state's students and culture is not discernible from these data.

Second, based on research by Boyd and colleagues, we hypothesized that the lack of familiarity effects will occur during an out-of-state prepared instructor's first year teaching, before they become acculturated to the state's educational environment and culture (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009). To test this explanation we limited our sample to first year teachers only and compared the effectiveness of out-of-state prepared instructors to in-state prepared and alternative entry teachers. Here, Table 2.6 shows that first year in-state prepared teachers are significantly more effective

in two mathematics models (rich covariate and school fixed effects) while alternative entry teachers are significantly more effective in one mathematics (rich covariate) and two reading models (school and school-by-year fixed effects). These alternative entry findings may be more indicative of familiarity with the state's culture—first year alternative entry teachers also lack exposure to the state's schools and curricula gained during traditional training—while the mathematics results for in-state prepared teachers may be attributable to greater familiarity with the educational context and/or culture. Overall, our two tests provide some support for the lack of familiarity hypothesis.

Teacher Turnover and Differential Attrition: To empirically test our final hypothesis we started by estimating the odds of teachers exiting North Carolina public schools. Here, results in Table 2.7 indicate that in-state prepared teachers have approximately one-half the odds of exiting North Carolina public schools as their out-of-state prepared peers, while alternative entry instructors are significantly more likely to exit teaching in the state. These results are consistent with our hypothesis that as out-of-state prepared teachers gain human capital through teaching experience in North Carolina, they will become more competitive for open teaching positions in their states of origin. These high rates of turnover for out-of-state prepared teachers could also signal a lack of commitment to teaching in the state that is manifested in a withdrawal of job-related effort. We further investigate this possibility in our differential attrition analyses.

In order to investigate potential differential attrition, we limited our sample to out-of-state prepared teachers only. Next, using NCDPI certified salary files, we created two attrition variables: (1) will leave, a time-invariant indicator equal to one for out-of-state prepared teachers who exited North Carolina public schools before beginning their sixth

year of teaching; and (2) last year, a time-varying indicator equal to one for out-of-state prepared teachers who did not return to North Carolina public schools in the following school year. Models controlling for will leave indicate whether more or less effective out-of-state prepared teachers exited North Carolina public schools, while models controlling for last year determine whether out-of-state prepared teachers not returning to North Carolina public schools the following school year are more or less effective than those who will stay for another year. Originally, we hypothesized that the differential attrition of the most effective out-of-state prepared teachers might explain their overall ineffectiveness. However, given the high rates of turnover for out-of-state prepared teachers shown in Table 2.7, results from this last year analysis are particularly important, because if out-of-state prepared teachers exit the state at substantially higher rates and are significantly less effective upon exiting (less effective than staying teachers or themselves in previous school years), this would provide evidence to understand their overall ineffectiveness.

Examining the top portion of Table 2.8, it is clear that the most effective out-of-state prepared teachers are not exiting North Carolina public schools. In the rich-covariate and school fixed effects specifications, out-of-state prepared teachers who will leave are significantly less effective than peers who will stay. This indicates that the differential attrition of the most effective out-of-state prepared teachers does not explain their overall effectiveness findings. Next, the middle portion of Table 2.8 demonstrates that out-of-state prepared teachers who will not return to North Carolina public schools the following school year are significantly less effective than their peers who will stay another year. While these last year results may simply indicate that exiting out-of-state

prepared teachers are less effective than peers who stay, it is also possible that exiting out-of-state prepared teachers knew they were going to leave North Carolina and that the last year findings are due to a withdrawal of job-related effort (Ashenfelter, 1978). To examine this possibility we included a teacher fixed effect—comparing effectiveness within teachers over time—in our last year models to determine whether out-of-state prepared teachers' effectiveness dips in their final year. Results from these models in the bottom panel of Table 2.8 indicate that in their last year out-of-state prepared teachers are significantly less effective in mathematics than in previous years; no evidence exists for effectiveness drops in elementary grades reading.

Due to the significant effects of teacher turnover on school district budgets and student achievement, the high rate of attrition for out-of-state prepared teachers is a concern (Alliance for Excellent Education, 2004; Ronfeldt, Loeb, & Wyckoff, 2013). While we originally hypothesized that the differential attrition of the most effective out-of-state prepared teachers might explain their underperformance, it is actually the differential attrition of less effective out-of-state prepared teachers—less effective than peers who remain and themselves in previous years (math only)—coupled with high turnover rates that help explain our findings.[10] Overall, it appears that out-of-state prepared teachers who leave North Carolina public schools are less effective, and since they exit in large numbers, their performance while in the state's classrooms brings down the average effectiveness of the preparation category as a whole.

[10] Relative to the results shown in Table 2.3, coefficients from models that compare out-of-state prepared teachers who will *stay* with in-state prepared and alternative entry instructors are reduced in size by approximately one-quarter (mathematics) and one-half (reading), respectively. These results, available from the authors upon request, provide further evidence that attrition patterns help explain the overall performance of out-of-state prepared teachers.

DISCUSSION

In response to a need for both more and better teachers, over the past two decades states have experimented with alternatives to solely licensing instructors prepared at in-state traditional education programs. While alternative entry programs have garnered the most policy and research attention during this time, many states have also reduced barriers to employment and broadened their potential labor pool of traditionally prepared instructors by granting reciprocal certification for out-of-state prepared teachers. Collectively, these reciprocal certification policies have helped facilitate the interstate movement of teachers and have aided high-growth states meet the demand for more teachers. Until recently, however, the effects of this policy choice were largely unexplored. Therefore, in this study we separated traditionally prepared teachers into two groups—in-state and out-of-state prepared—and assessed both the comparative effectiveness of out-of-state prepared instructors and potential explanations for differences in out-of-state prepared teacher performance.

Results indicated clear effectiveness differences within the traditionally prepared teacher population: in-state prepared teachers significantly outperformed their out-of-state prepared peers in elementary school mathematics and reading across all model specifications. Furthermore, alternative entry teachers were also significantly more effective across all subjects and models. These alternative entry findings are also fairly robust to model specifications that separate two high performing groups of teachers—Teach For America and Visiting International Faculty—from the main alternative entry category. Overall, out-of-state prepared teachers are the second largest and least effective source of early-career instructors in North Carolina elementary schools. In response to

these findings, other states may wish to perform similar analyses to assess the effects of their own reciprocal certification policies.

Moving forward with effective policy responses requires an understanding of why out-of-state prepared teachers struggle, and here, our results suggest two factors that help explain out-of-state prepared teachers' performance. First, out-of-state prepared teachers are less effective due to their lack of familiarity with North Carolina's educational environment and culture. Out-of-state prepared teachers trained in contiguous states perform comparably to in-state prepared teachers in reading while those trained in non-contiguous states are less effective; first year in-state prepared teachers are more effective than first year out-of-state prepared instructors in mathematics while first year alternative entry teachers are more effective in reading. These findings are consistent with the composition of imported teachers in North Carolina—a majority come from more urban, distant regions of the Northeast and Midwest, particularly New York, Pennsylvania, Michigan, and Ohio. Furthermore, these findings align with previous research which indicates the importance of training/preparation experiences that mirror the educational environment of future employment (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009). Second, early-career out-of-state prepared teachers are less effective due to their attrition patterns. The odds that out-of-state prepared teachers will exit North Carolina public schools are approximately twice those of in-state traditionally prepared instructors—this attrition may be attributable to teachers seeking positions in their states of origin and/or difficulties faced by out-of-state prepared teachers from more distant states assimilating into the communities where they work. This turnover, coupled with the ineffectiveness of departing out-of-state prepared teachers (less effective than peers

who stay and themselves in previous years), indicates that the average effectiveness of out-of-state teachers is brought down by those who leave. Overall, many North Carolina students are taught by less effective, exiting out-of-state prepared instructors.

In response to these findings several policy solutions seem appropriate. First, North Carolina could increase the total number of in-state prepared teachers in the workforce by improving the yield for in-state preparation programs—the percentage of graduates of in-state preparation programs hired in North Carolina. Second, North Carolina could also increase the production of in-state prepared (undergraduate, graduate, and licensure only) teachers and allow more alternative entry instructors into the workforce. There are two potential concerns with these policy solutions: (1) increasing the quantity of newly prepared in-state and alternative entry teachers could compromise the quality of those teachers if selection/hiring requirements are lowered or if in-state institutions of higher education reduce the quality of preparation in response to growth; and (2) due to the increasing student population of North Carolina public schools, the in-state and alternative entry supply is unlikely to meet demand in the short-term. As alternative policy mechanisms the state could: (1) institute rigorous selection/hiring practices for out-of-state prepared teachers, including intensive recruiting in contiguous states (or those states deemed similar to North Carolina) and focusing on teachers' non-cognitive characteristics (Duckworth, Quinn, & Seligman, 2009; Rockoff, Jacob, Kane, & Staiger, 2011); and (2) direct increased resources to induction, mentorship, and other support services that ease an out-of-state prepared teacher's transition to the state's curriculum and work environment and aid their assimilation into the communities in which they teach. Finally, the attrition findings from this work—exiting out-of-state

prepared teachers adversely impact student achievement—suggest that in addition to better selecting teachers and promoting on-the-job learning with novice teacher induction programs, it may be beneficial for states, assuming the challenges of using teacher value-added scores can be overcome, to proactively filter out less effective early-career teachers.

Given the current labor market context, where individuals are more mobile and change positions more frequently, understanding the potential effects of greater mobility for the teaching profession is critical. Our work suggests that teachers prepared in other states address areas of shortage in importing states, but, on average, are not familiar enough with or committed enough to the importing state. Therefore, we contend that states must craft policies that recognize the labor market(s) for their own state in order to improve teacher effectiveness.

REFERENCES

Alliance for Excellent Education. (2004). Tapping the potential: Retaining and developing high-quality teachers. Available from: http://www.all4ed.org/publication_material/reports/tapping_potential.

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *Review of Economics and Statistics, 60*(1), 47-57.

Bacolod, M. (2007). Do alternative opportunities matter? The role of female labor markets in the decline of teacher quality. *The Review of Economics and Statistics, 89*(4), 737-751.

Bifulco, R. (2012). Can non-experimental estimates replicate estimates based on random assignment in evaluations of school choice? A within-study comparison. *Journal of Policy Analysis and Management, 31*(3), 729-751.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). The draw of home: How teachers' preferences for proximity disadvantage urban schools. *Journal of Policy Analysis and Management, 24*(1), 113-132.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & 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., Goldhaber, D., Lankford, H., & Wyckoff, J. (2007). The effect of certification and preparation on teacher quality. *The Future of Children, 17*(1), 45-68.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis, 31*(4), 416-440.

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis, 24*(4), 305-331.

Chingos, M., & Peterson, P. (2011). It's easier to pick a good teacher than to train one: Familiar and new results on the correlates of teacher effectiveness. *Economics of Education Review, 30*(3), 449-465.

Clotfelter, C., Ladd, H., & Vigdor, J. (2007.) Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682.

Clotfelter, C., Ladd, H., & Vigdor, J. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *The Journal of Human Resources, 45*(3), 655-6861.

Common Core of Data (CCD), State non-fiscal survey of public elementary/secondary education. 2000-01 v.1c, 2001-02 v.1c, 2002-2003 v.1b, 2003-04 v.1b, 2004-05 v.1f, 2005-06 v.1b, 2006-07 v.1c, 2007-08 v.1b, 2008-09 v.1c, 2009-10 v.1c.

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

Dobbie, W. (2011). Teacher characteristics and student achievement: Evidence from Teach For America. Available from: http://www.people.fas.harvard.edu/~dobbie/research/TeacherCharacteristics_July2011.pdf.

Duckworth, A., Quinn, P., & Seligman, M. (2009). Positive predictors of teacher effectiveness. *The Journal of Positive Psychology, 4*(6), 540-547.

Feistritzer, C.E. (2011). *Profiles of teachers in the U.S. 2011*. Washington, D.C.: National Center for Education Information (downloaded on July 29, 2011, from http://www.edweek.org/media/pot2011final-blog.pdf).

Glazerman, S., Levy, D., & Myers, D. (2003). Non-experimental versus experimental estimates of earnings impacts. *The ANNALS of the American Academy of Political and Social Science, 589*(1), 63-93.

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.

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

Goldhaber, D., & Liddle, S. (2011). The Gateway to the Profession: Assessing Teacher Preparation Programs Based on Student Achievement. *Center for Education Data and Research.*

Goldhaber, D., Gross, B, & Player, D. (2011). Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best? *Journal of Policy Analysis and Management, 30*(1), 57-87.

Greenberg, J., Pomerance, L., & Walsh, K. (2011). Student teaching in the United States. *National Council on Teacher Quality.* Available from: www.nctq.org/edschoolreports/studentteaching/.

Greenwald, R., Hedges, L., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Educational Research, 66(3),* 361-396.

Henry, G.T., Bastian, K.C., & Fortner, C.K. (2011). Stayers and Leavers: Early-career teacher effectiveness and attrition. *Educational Researcher, 40*(6), 271-280.

Henry, G.T., Thompson, C.L., Fortner, C.K., Bastian, K.C., & Marcus, J.V. (2011). Technical report: UNC teacher preparation program effectiveness. *Carolina Institute for Public Policy.* Available from: http://publicpolicy.unc.edu/research/publications-presentations-and-reports.

Henry, G.T. Fortner, C.K., & Bastian, K.C. (2012). The effects of experience and attrition for novice high-school science and mathematics teachers. *Science, 335*, 1118-1121.

Henry, G.T., Bastian, K.C. & Smith, A.A. (2012). Scholarships to recruit the "best and brightest" into teaching: Who is selected, where do they teach, how effective are they, and who long do they stay. *Educational Researcher, 41*(3), 83-92.

Henry, G.T., Bastian, K.C., Fortner, C.K., Kershaw, D.C., Purtell, K.M., Thompson, C.L., & Zulli, R.A. (2013). Teacher preparation policies and their effects on student achievement. In press, *Education Finance and Policy.*

Ingersoll, R., & Smith, T. (2003). The wrong solution to the teacher shortage. *Educational Leadership, 60*(8), 30-33.

Ingersoll, R., & Merrill, L. (2010). Who's teaching our children? *Educational Leadership, 67*(8), 14-21.

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

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

National Research Council. (2010). *Preparing teachers: Building evidence for sound policy*. Washington, DC: National Academies Press.

North Carolina Department of Public Instruction, *North Carolina Mathematics Tests. Ed. 3: Technical Report*, Raleigh, NC, downloaded December 29, 2011 from (http://www.dpi.state.nc.us/docs/accountability/reports/mathtechmanualdrafted2.pdf).

Porter, A.C., Polikoff, M.S., & Smithson, J. (2009). Is there a de facto national intended curriculum? Evidence from state content standards. *Educational Evaluation and Policy Analysis, 31*(3), 238-268.

Reininger, M. (2012). Hometown disadvantage? It depends on where you're from: Teacher's location preferences and the implications for staffing schools. *Educational Evaluation and Policy Analysis, 34*(2), 127-145.

Rockoff, J., Jacob, B., Kane, T., & Staiger, D. (2011). Can you recognize an effective teacher when you recruit one? *Education Finance and Policy, 6*(1), 43-74.

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal, 50*(1), 4-36.

Shadish, W., Clark, M., & Steiner, P. (2010). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association, 103*(484), 1334-1343.

Shen, J. (1997). Has the alternative certification policy materialized its promise? A comparison between traditionally and alternatively certified teachers in public schools. *Educational Evaluation and Policy Analysis, 19*(3), 276-283.

Shrout, P.E., & Bolger, N. (2002). Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psychological Methods, 7*(4), 422-445.

Table 2.1: Covariates for Analyses

Student Covariates	Classroom and Teacher Covariates	School Covariates
1) Prior test scores (reading and math 2) Classmates prior test scores 3) Days absent 4) Structural mobility 5) Within year mobility 6) Other between year mobility 7) Race/ethnicity 8) Gender 9) Poverty status 10) Gifted 11) Disability 12) Currently limited English proficient 13) Was limited English proficient 14) Overage for grade 15) Underage for grade	16) Class size 17) Heterogeneity of prior achievement within the classroom 18) Teacher out-of-field status 19) Single year indicators for teacher experience 20) Teacher preparation categories a. Out of state prepared (reference group) b. In-state prepared c. Alternative entry	21) School size 22) School size squared 23) Violent acts per 1,000 students 24) Short-term suspension rate 25) Total per-pupil expenditures 26) District teacher supplements 27) Racial/ethnic composition 28) Concentration of poverty

Note: We included these covariates in our preferred rich covariate adjustment and school fixed effects analyses. Student fixed effects exclude time-invariant student characteristics; school-year fixed effects exclude school-level variables. Most models testing hypotheses for out-of-state prepared teacher ineffectiveness include these covariates, substituting or adding variables as needed per the empirical test performed.

Table 2.2: Teacher Preparation Category Descriptive Information

	Students' Average Prior EOG Scores		Classroom Percent Free or Reduced Price Lunch		Average School Performance Composite		School Percent Free or Reduced Price Lunch	
	ES Math	ES Read	ES Math	ES Read	ES Math	ES Read	ES Math	ES Read
Out-of-State Prepared	-0.069 (0.539)	-0.079 (0.510)	47.77 (26.59)	49.27 (26.29)	60.39 (16.39)	60.38 (16.35)	54.48 (25.81)	54.63 (25.80)
In-State Prepared	-0.073 (0.482)	**-0.069*** (0.465)	**50.60**** (25.07)	**51.42**** (24.44)	60.48 (15.27)	60.56 (15.23)	**57.88**** (23.58)	**57.82**** (23.50)
Alternative Entry	**-0.335**** (0.602)	**-0.343**** (0.611)	**53.32**** (28.75)	**55.42**** (28.04)	**54.57**** (16.48)	**54.67**** (16.44)	**65.53**** (25.44)	**65.52**** (25.35)

Note: In the table above, students' average prior EOG scores and classroom percent free reduced-price lunch identify unique teacher-classroom combinations. Average school performance composite (number of tests passed at a school divided by the number of tests taken) and school free reduced-price lunch identify unique teacher-school-year combinations. * Indicates values statistically different than those for out-of-state prepared teachers at the 0.05 level; ** indicates values statistically different than those for out-of-state prepared teachers at the 0.01 level.

Table 2.3: Elementary School Mathematics and Reading Outcomes

	Rich Covariate Adjustment		School Fixed Effects		School-Year Fixed Effects		Student Fixed Effects	
Teacher Prep Category	**ES Math**	**ES Reading**	**ES Math**	**ES Reading**	**ES Math**	**ES Reading**	**ES Math**	**ES Reading**
In-state prepared	**0.024**** (0.004)	**0.012**** (0.003)	**0.019**** (0.004)	**0.011**** (0.003)	**0.019**** (0.004)	**0.009**** (0.003)	**0.023**** (0.005)	**0.008*** (0.004)
Alternative entry	**0.030**** (0.007)	**0.013*** (0.005)	**0.029**** (0.007)	**0.025**** (0.005)	**0.021**** (0.008)	**0.025**** (0.006)	**0.019*** (0.009)	**0.017**** (0.006)
Observations Used	715,172	1,008,362	638,290	883,837	492,354	673,522	120,677	170,691
Student Covariates	✓	✓	✓	✓	✓	✓	✓	✓
Classroom/Teacher Covariates	✓	✓	✓	✓	✓	✓	✓	✓
School Covariates	✓	✓	✓	✓			✓	✓
Unique Teacher Counts Contributing to Preparation Category Estimates								
Out-of-state prepared	4,197	4,255	4,130	4,195	3,820	3,877	3,544	3,620
In-state prepared	5,330	5,413	4,509	4,603	3,672	3,753	3,577	3,682
Alternative entry	1,278	1,323	1,083	1,117	871	903	753	796

Note: All coefficients are in relation to out-of-state prepared teachers. * indicates significance at the 0.05 level; ** indicates significance at the 0.01 level.

Table 2.4: Examining the Quality of Imported Instructors

	Do Teacher Test Scores Significantly Predict Student Achievement Gains?					
	Rich Covariate Adjustment		**School Fixed Effects**		**School-Year Fixed Effects**	
	ES Math	**ES Read**	**ES Math**	**ES Read**	**ES Math**	**ES Read**
Standardized Teacher Test Scores	**0.029**** (0.003)	**0.005*** (0.002)	**0.028**** (0.003)	**0.005*** (0.002)	**0.026**** (0.003)	0.005 (0.003)
Observations Used	715,172	1,008,362	714,198	1,006,001	679,625	954,991
	Do Teacher Test Scores Mediate the Teacher Preparation Effects?					
	Rich Covariate Adjustment		**School Fixed Effects**		**School-Year Fixed Effects**	
	ES Math	**ES Read**	**ES Math**	**ES Read**	**ES Math**	**ES Read**
In-State Prepared	**0.023**** (0.004)	**0.012**** (0.003)	**0.019**** (0.004)	**0.011**** (0.003)	**0.018**** (0.004)	**0.009**** (0.003)
Alternative Entry	**0.035**** (0.007)	**0.013*** (0.005)	**0.034**** (0.007)	**0.026**** (0.005)	**0.025**** (0.008)	**0.026**** (0.006)
Standardized Teacher Test Scores	**0.030**** (0.003)	**0.005*** (0.002)	**0.029**** (0.003)	**0.006*** (0.003)	**0.027**** (0.003)	**0.006*** (0.003)
Observations Used	715,172	1,008,362	638,290	883,837	492,354	673,522

Note: The coefficients on standardized teacher test scores in the top panel test whether this measure of human capital significantly predicts student achievement. The coefficients in the bottom panel, when compared to those in Table 2.3, indicate whether teacher test scores mediate the differences in teacher preparation category effectiveness. * indicates significance at the 0.05 level; ** indicates significance at the 0.01 level.

Table 2.5: Out-of-State Teacher Performance from Contiguous and Non-Contiguous States

Teacher Preparation Category	Out-of-State Teachers from Contiguous States as the Reference Group					
	Rich Covariate Adjustment		School Fixed Effects		School-Year Fixed Effects	
	ES Math	ES Read	ES Math	ES Read	ES Math	ES Read
In-state prepared	**0.021**** (0.008)	0.005 (0.006)	**0.032**** (0.008)	0.007 (0.006)	**0.032**** (0.009)	0.004 (0.006)
Alternative entry	**0.027**** (0.010)	0.006 (0.007)	**0.041**** (0.010)	**0.019**** (0.007)	**0.034**** (0.011)	**0.019*** (0.008)
Observations Used	490,941	702,387	206,935	287,201	111,911	152,705
Teacher Preparation Category	**Out-of-State Teachers from Non-Contiguous States as the Reference Group**					
	Rich Covariate Adjustment		**School Fixed Effects**		**School-Year Fixed Effects**	
	ES Math	**ES Read**	**ES Math**	**ES Read**	**ES Math**	**ES Read**
In-state prepared	**0.025**** (0.004)	**0.013**** (0.003)	**0.018**** (0.004)	**0.013**** (0.003)	**0.018**** (0.004)	**0.011**** (0.003)
Alternative entry	**0.031**** (0.007)	**0.014*** (0.005)	**0.030**** (0.007)	**0.026**** (0.005)	**0.021**** (0.008)	**0.027**** (0.006)
Observations Used	680,568	961,225	582,234	799,991	434,349	592,083

Note: All coefficients are in relation to out-of-state prepared teachers (contiguous or non-contiguous state preparation). * indicates significance at the 0.05 level; ** indicates significance at the 0.01 level.

Table 2.6: Comparing the Effectiveness of First Year Teachers

Teacher Preparation Category	Rich Covariate Adjustment		School Fixed Effects		School-Year Fixed Effects	
	ES Math	ES Reading	ES Math	ES Reading	ES Math	ES Reading
In-state prepared	**0.023**** (0.008)	0.006 (0.006)	**0.023**** (0.009)	0.003 (0.007)	0.010 (0.013)	0.005 (0.010)
Alternative entry	**0.035*** (0.015)	0.020 (0.12)	0.028 (0.015)	**0.029*** (0.013)	0.017 (0.022)	**0.052**** (0.018)
Observations Used	147,560	207,437	101,305	138,953	46,872	63,759

Note: All coefficients are in relation to first-year out-of-state prepared teachers. * indicates significance at the 0.05 level; ** indicates significance at the 0.01 level.

Table 2.7: Logistic Regression Results for Teacher Turnover

Teacher Category	Odds Ratio and Z-Score
In-State Prepared	**0.464**** (-14.48)
Alternative Entry	**1.291**** (3.62)
Observations Used	22,910

Note: Odds ratios are in relation to out-of-state prepared teachers, with teachers exiting North Carolina public schools as the dependent variable. *indicates significance at the 0.05 level;
** indicates significance at the 0.01 level.

Table 2.8: Teacher Effectiveness for Out-of-State Teachers Who Stay or Leave

Focal Variables	Rich Covariate Adjustment		School Fixed Effects		School-Year Fixed Effects	
	ES Math	ES Read	ES Math	ES Read	ES Math	ES Read
Will Leave	**-0.024**** (0.007)	**-0.024**** (0.005)	**-0.021**** (0.007)	**-0.020**** (0.005)	-0.001 (0.008)	-0.011 (0.006)
Observations Used	260,144	354,849	188,603	254,911	109,967	147,039
Last Year	**-0.035**** (0.008)	**-0.025**** (0.007)	**-0.033**** (0.008)	**-0.024**** (0.006)	-0.017 (0.010)	**-0.019*** (0.008)
Observations Used	260,144	354,849	183,146	248,900	80,369	107,890

Teacher Fixed Effects: Testing for an Ashenfelter Dip		
Focal Variable	**ES Math**	**ES Read**
Last Year	**-0.021*** (0.010)	-0.000 (0.009)
Observations Used	32,386	43,503

Note: Coefficients for the top two models are in relation to out-of-state prepared teachers who will not exit North Carolina public schools before beginning a sixth year of teaching or who will return for another school year. In the teacher fixed effects model an out-of-state prepared teacher's effectiveness in his/her last year is compared with his/her effectiveness in previous school years. *indicates significance at the 0.05 level; ** indicates significance at the 0.01 level.

CHAPTER THREE

The Apprentice: Pathways to the Principalship and Their Effects on Student Achievement[1]

SUMMARY

Given the high concentration of early-career principals in the workforce and the adverse academic effects in schools transitioning to a first-time principal, we detail the characteristics of first-time principals and examine the relationship between characteristics of early-career principals and the environments in which they previously worked and student achievement gains. Descriptively, we find that first-time principals are "homegrown"—hired from within the district—and evidence that first-time principals sort into schools based in part on their individual characteristics and professional credentials. Value-added analyses indicate that several individual principal characteristics are significantly associated with student achievement gains; however, our strongest findings concern characteristics of prior work environments. Specifically, results suggest that early-career principals who served in high value-added schools as assistant principals are more effective principals. Further research must more rigorously test hypotheses from our value-added findings and develop policy recommendations.

[1] Kevin C. Bastian is the first author for this research and was responsible for the data management, analysis, organization and writing of the paper. Gary T. Henry is the second author and contributed to the research focus, framing, and editing of the paper.

INTRODUCTION

A principal's leadership can influence whether students, teachers, and the school as a whole succeed. Therefore, research attention is increasingly focused on principals. Key research issues center on (1) the magnitude of principal effects on outcomes of interest—student achievement, absences, and graduation rates; teacher retention and on-the-job learning; and school working conditions (Boyd, Grossman, Ing, Lankford, Loeb, & Wyckoff, 2011; Branch, Hanushek, & Rivkin, 2012; Coelli & Green, 2012; Dhuey & Smith, 2012; Loeb, Kalogrides, & Beteille, 2012); (2) the relationship between principal actions, such as recruiting and retaining high quality teachers or acting as the school's instructional leader, and principal effectiveness (Grissom & Loeb, 2011; Horng, Klasik, & Loeb, 2010; Robinson, Lloyd, & Rowe, 2008); (3) the characteristics of principals—demographics, career trajectories, preparation—and; (4) whether principal characteristics, such as experience or preparation type, influence their effectiveness (Clark, Martorell, & Rockoff, 2009; Corcoran, Schwartz, & Weinstein, 2012).

While these issues are relevant for all principals, two factors suggest that they are particularly pressing for inexperienced principals. First, inexperienced principals comprise a significant percentage of school leaders. For example, as shown in Figure 3.1, during the most recent five years of available data (2006-07 through 2010-11), first-time principals (those with no prior experience in this position) made up 10.83 percent of North Carolina's principal workforce—the third highest percentage—and overall, principals with less than five years of experience (early-career principals) comprised 54 percent of all public school leaders in the state (authors' analysis).

Second, evidence suggests that a school's transition to a first-time principal adversely impacts student achievement (Burkhauser, Gates, Hamilton, & Ikemoto, 2012). Using the present study's sample, Figure 3.2 shows a sharp drop in standardized school effectiveness (school value-added) during the transition to a first-time principal in four of five grade-level/subject comparisons—elementary grades mathematics and reading, middle grades reading, and high school. On average, these effectiveness decreases were approximately ten percent of a standard deviation, and while school performance generally improved in a principal's second year (except for elementary grades reading), values were still below the pre-transition performance, particularly in elementary schools.

Therefore, given the prevalence of inexperienced principals and adverse achievement effects for schools transitioning to a first-time principal, the present study seeks to contribute to the nascent principal research agenda by better understanding pathways to the principalship and whether characteristics of principals or those of their prior work environments are associated with principal effectiveness. Specifically, we ask the following:

(1) What are the characteristics of first-time principals, including their prior work experiences in educational settings, and the schools that hire them?

(2) Are there characteristics of early-career principals or the environments in which they previously worked that are associated with student achievement gains?

By way of preview, we find that first-time principals are "homegrown"—a large majority become first-time principals within the same district in which they once worked as teachers and/or assistant principals—and evidence of first-time principals sorting into schools based on their individual characteristics and professional credentials. Concerning

student achievement gains, we must caution that our estimates are not causal, but rather, identify associations of interest that serve as the foundation for hypothesis generation and future, more methodologically rigorous research. Overall, several individual principal characteristics were significantly associated with student achievement, however, our strongest findings concerned characteristics of principals' prior work environments, especially a congruence between the level (middle/high) of the school where they served as assistant principals and the schools where they began as principals and the effectiveness (value-added) of their assistant principal placement school. Both of these findings suggest that first-time principals learn key aspects of how to be an effective school leader during their assistant principal experience.

In the remainder of this article we first provide an overview of the methodological challenges in estimating principal effectiveness and research regarding influences on principal effectiveness. Next, we detail our data sources, research sample, and analysis measures. Then, we describe our methodological approach and threats to estimate validity. Finally, we present results for each research question and conclude with a discussion of future research directions and potential policy implications.

BACKGROUND

Estimating Principal Effectiveness

Recognizing the effects of principals on student achievement, researchers are beginning to employ rich administrative datasets and sophisticated econometric techniques to investigate the magnitude and determinants of principal effectiveness. While promising, this work presents several conceptual and methodological challenges that we review below.

First, researchers are primarily defining principal effectiveness or examining the effects of particular principal characteristics (e.g. preparation type, years of experience) based on the education production function and value-added to student achievement.[2] Unlike teachers, however, principals do not directly affect student learning, and instead, researchers must assume that measured academic outcomes attributed to principals are due to their indirect effects—through mechanisms such as hiring and retaining teachers, establishing a school culture, improving instructional quality, or managing school resources (Branch, Hanushek, & Rivkin, 2012). Second, in addition to indirectly influencing outcomes, principals' effects may not be immediate, but rather, may develop over time. For example, research suggests that the effects of principals accrue over several years, perhaps as they are able to affect the human capital of their teachers and shape school culture (Coelli & Green, 2012; Handa, Thompson, Marcus, & Smith, 2010). Third, when estimating principal effects researchers need to: (1) be aware of endogeneity threats caused by principals sorting into schools based on (un)observed characteristics and (2) carefully consider what variables to control for in an education production function (Beteille, Kalogrides, & Loeb, 2012; Clark, Martorell, & Rockoff, 2009; Loeb, Kalogrides, & Horng, 2010). For instance, whether it is appropriate to control for elements of teacher quality (e.g. type of preparation) at a school likely depends on the amount of autonomy district policies or labor markets provide to principals to influence the human capital of their teacher workforce (Clark, Martorell, & Rockoff, 2009). This question of principal control highlights a final conceptual challenge: whether and how to

[2] Recently, research articles have primarily focused on student achievement gains; some have also examined principal effects on student absences, student behavior outcomes, student graduation rates, teacher retention, and teacher on-the-job effectiveness gains (Branch, Hanushek, & Rivkin, 2012; Clark, Martorell, & Rockoff, 2009; Coelli & Green, 2012; Loeb, Kalogrides, & Beteille, 2012).

separate the effects of principals from other inputs of a school—e.g. the school's performance trajectory, financial resources, neighborhood composition, or unmeasured school characteristics.[3]

To address threats to estimate validity or to separate principal from school effects, researchers often specify an education production function with fixed effects. Here: (1) principal fixed effects attempt to control for time-invariant principal characteristics, such as ability, and limit effectiveness comparisons to within principals over time (Branch, Hanushek, & Rivkin, 2012); (2) school fixed effects attempt to separate principal effects from school effects by comparing different principals working in the same schools over time (Branch, Hanushek, & Rivkin, 2012; Clark, Martorell, & Rockoff, 2009); and (3) principal and school fixed effects (combined) attempt to control for unobserved principal characteristics and school factors by limiting comparisons to principals who change schools (Coelli & Green, 2012; Dhuey & Smith, 2012). While these approaches preference internal validity, sampling bias may be an issue due to limitations on the samples for which the effects are identified—principals that change schools and schools that change principals may be quite different from the full population (Chiang, Lipscomb, & Gill, 2012). Finally, depending upon the theoretical model of principal effectiveness—that is, whether the mechanisms through which principals are expected to affect an outcome of interest are hypothesized to develop quickly or slowly—researchers can specify models to identify immediate effects or longer-term school/principal improvement (Grissom, Kalogrides, & Loeb, 2012).

[3] While researchers have preferred to identify principal effects (separate from those of the school), recent research evinces that school effectiveness estimates (not principal effectiveness) better align with non-test-based measures of principal quality (Grissom, Kalogrides, & Loeb, 2012).

Research on Principal Effectiveness and Its Determinants

In studies using the methods described above, recent findings indicate that principals significantly influence student academic outcomes. For principals switching schools in Texas, a one standard deviation increase in principal effectiveness is associated with a 0.10 standard deviation increase in student achievement (Branch, Hanushek, & Rivkin, 2012); for principals switching schools in British Columbia, a one standard deviation increase in principal effectiveness boosts student achievement by approximately 0.10 to 0.20 standard deviations between 4^{th} and 7^{th} grade exams (Dhuey & Smith, 2012). Additionally, a principal one standard deviation higher in the effectiveness distribution increases graduation rates by 2.6 percentage points—roughly one-third of the standard deviation across schools (Coelli & Green, 2012)—and principals in the top quartile of effectiveness are associated with significantly higher student attendance rates (Branch, Hanushek, & Rivkin, 2012).

To understand what factors account for these performance variations researchers have begun examining the effects of three principal characteristics—principal preparation, principal experience or tenure at a school, and principal career trajectories. Regarding principal preparation, results indicate that: (1) principals holding a Master's of School Administration from a University of North Carolina system institution are no more or less effective than principals with preparation from private universities or from out-of-state (Handa, Thompson, Marcus, & Smith, 2010); (2) principals prepared in university training programs are no more effective than those prepared in school district programs (Vanderhaar, Munoz, & Rodosky, 2006); (3) schools led by graduates of New York City's *Aspiring Principals Program* are initially lower-performing but narrow gaps

over time (Clark, Martorell, & Rockoff, 2009; Corcoran, Schwartz, & Weinstein, 2012); and (4) graduates of the *New Leaders for New Schools* program are less effective in their first year at a school, but more effective than comparison principals in subsequent years (Martorell, Heaton, Gates, & Hamilton, 2010). Overall, type of principal preparation does not appear to have a sizable effect on student achievement, but as more alternative preparation programs develop and university-based programs revamp their curriculum, type and components of principal preparation merit continued research.[4]

Concerning returns to principal experience, results are mixed across studies. Findings from some early and more recent research indicate no relationship between principal experience and effectiveness (Brewer, 1993; Dhuey & Smith, 2012). Other work, especially studies using single-year principal experience indicators, display significant early-career principal returns to experience (Branch, Hanushek, & Rivkin, 2012; Clark, Martorell, & Rockoff, 2009; Eberts & Stone, 1988). Regardless of overall principal experience, school academic outcomes also appear to improve with principal tenure at a school, indicating that it may take time for principals to substantially influence student performance (Handa, Thompson, Marcus, & Smith, 2010). Finally, only one study to date has examined variables related to work experiences prior to becoming a principal—whether an individual was ever a teacher in the same school or an assistant principal in the same school in which he/she became the principal. Here, evidence indicates that for principals in their first two years ever serving as an assistant principal in

[4] As an example of university-based programs revamping their principal preparation curriculum, North Carolina recently opened three Regional Leadership Academies with federal Race to the Top and Z. Smith Reynolds funding. Participants in these Leadership Academies are recruited from school district nominees in each region, earn credits towards a Master's in School Administration through a case-study curriculum, have a full-time, year-long clinical residency experience in the high-need districts and schools they will serve as principals, and upon graduation are granted a North Carolina principal license.

the current principalship school is positively associated with student achievement gains; conversely, ever serving as a teacher in the current principalship school is negatively associated with student achievement gains (Clark, Martorell, & Rockoff, 2009).

Overall, the research agenda concerning influences on principal effectiveness remains in nascent stages, with much to learn about who becomes a principal and how the individual characteristics of principals or the environments in which they have previously worked influence student achievement. Below, we explicate how a principal's prior work environments, particularly the assistant principalship, may influence early-career principal effectiveness.

Why Prior Work Environments May Matter

In contrast to the traditional pathway into teaching—a four-year education degree, student teaching, and an immediate transition to in-service teaching—promotion to the principalship often occurs several years after completing the degree requirements for a principal license. Generally (as shown in Table 3.2), principals begin their careers as teachers, acquire a principal license through formal education (sometimes while still teaching), and serve an apprenticeship of varying length as an assistant principal. During this progression, especially the apprenticeship, individuals likely experience a tremendous amount of on-the-job learning that influences their future effectiveness as principals.

One lens through which to conceptualize this on-the-job learning is Bandura's self-efficacy model. This postulates that individuals' efficacy expectations are most malleable early in learning—especially salient to assistant principals and early-career principals—and are shaped by mastery and vicarious experiences, attribution beliefs, and

social persuasion (Bandura, 1977).[5] For assistant principals, mastery experiences would come through directly engaging in practices, such as formally observing teachers, participating in teacher hiring, or analyzing student achievement data, to increase teacher and student performance; vicarious experiences would occur as sitting principals model effective school leadership behaviors. As apprentices witness these practices succeeding, the attribution of desired school outcomes to school leadership actions would be enhanced.

To examine whether prior learning environments influence early-career principal performance, we consider several measures. These include: (1) previously serving as a teacher or assistant principal in the principalship school, which may suggest knowledge of the particular environment that promotes success; (2) working as an assistant principal at the same school level (elementary, middle, or high) as the principalship school, which my familiarize a principal with the requirements of managing students, teachers, and school operations at that level; and (3) apprenticing in a high value-added school, which may suggest that assistant principals (future principals) are better prepared to implement effective strategies after directly participating in and/or being vicariously exposed to such practices themselves. Overall, the time between formal principal preparation and the assumption of school leadership suggests that other learning experiences, especially the apprenticeship learning environment, may shape early-career principal performance.

[5] The concept of efficacy is particularly relevant since research over the past two decades has found that teacher self-efficacy and the collective efficacy of schools is significantly associated with student achievement (Goddard, Hoy, & Woolfolk Hoy, 2000; Pajares, 1996).

DATA AND SAMPLE

Data Sources

In order to track the career trajectories and identify the individual characteristics of first-time principals in our research sample and create focal variables for our value-added analyses we relied on data from the University of North Carolina General Administration (UNCGA) and the North Carolina Department of Public Instruction (NCDPI). Specifically, the UNCGA supplied identifiers for graduates of in-state public university Masters of School Administration (MSA) programs and North Carolina Principal Fellows—a competitive, merit-based scholarship program offered at 11 in-state public institutions which culminates in a MSA degree. The NCDPI provided a comprehensive set of personnel data, including the following key files: (1) a certified personnel pay history file that allowed us to determine what positions (e.g. teacher, assistant principal) individuals held, how long they held those positions, in what districts and schools they worked, and when they first became a principal; (2) a personnel education file that detailed the level of degree individuals earned, the degree-granting institution, and the graduation date; (3) personnel licensure files that indicated individuals' teaching license areas and National Board Certification status; and (4) test score files that contained teacher and principal licensure exam scores.[6] In addition to these files, the NCDPI also provided student test score data, classroom rosters (to allow

[6] During the time period of this study, North Carolina was one of only 17 states (including the District of Columbia) that required principals to pass the Educational Testing Service (ETS) School Leadership Series licensure exams. In our sample of 981 first-time principals, most principals had taken the ETS Praxis 1010 exam, which is a six hour test that requires prospective principals to evaluate actions by responding to 16 vignettes, demonstrate problem solving with two case studies, and respond to a series of documents (assessment data, school improvement plans, budgets, etc.) to display analysis and decision-making skills.

matching of students and teachers), and school-level characteristics to facilitate our value-added analyses. More about these data is included in the measures section below.

Research Sample

For this analysis we were interested in identifying a set of first-time principals in North Carolina public schools and then (1) detailing their career trajectories and (2) examining whether individual characteristics or those of the environments in which they had previously worked are associated with student achievement gains. To define a principal-school combination in administrative data we specified the following rules: (1) an individual had to work 100 percent full-time equivalency as a principal at a school; (2) an individual had to begin work as a principal at a school in one of the fiscal year's first three pay periods (July, August, or September); and (3) after beginning work in one of the first three pay periods, an individual had to remain as a full-time principal for at least eight pay periods. Using these guidelines we created a research sample that consists of all first-time principals, 981 in total, in the 2006-07, 2007-08, 2008-09, or 2009-10 school years. Data described above in the data sources section allows us to identify the career trajectories (previous teaching and assistant principal work) and principal preparation of our sample. Additionally, student test score data to facilitate value-added analyses are available from 2004-05 through 2010-11. This means for our first cohort of first-time principals—starting in the 2006-07 school year—we have up to five years of data to examine their contributions to student achievement gains; for our last cohort of first-time principals—starting in the 2009-10 school year—we have up to two years of data. To both account for this unbalanced panel and examine whether the associations between student achievement gains and principal characteristics differ according to principal

experience, specification checks for our value-added models are limited to principals in their first two years and then to principals in their third, fourth, or fifth year.

Measures

Dependent Variables: To examine the associations between individual principal characteristics or those of their previous working environments and student achievement, the dependent variable for this analysis is students' current test score performance on either the North Carolina End-of-Grade (EOG) mathematics and reading exams (grades 3-8) or the five North Carolina high school (grades 9-12) End-of-Course (EOC) exams required for graduation (English 1, algebra 1, biology, civics, and U.S. history). We standardized all EOG exams within subject, grade, and year and all EOC exams within subject and year. To further control for secular trends or other year-to-year anomalies we include year fixed effects in our value-added specifications. Finally, standardized mathematics and reading scores from the previous grade, or from eighth grade for high school students, serve as the measure of prior achievement in these value-added models. More on our additional covariates is included in the sections below.

Focal Variables: For this analysis we examine the value-added associations between (1) individual principal characteristics or (2) characteristics of the environments in which early-career principals previously worked and student achievement gains. By conceptualizing the determinants of principal effectiveness in these two broad categories, results from this work can provide a clear policy focus for future research: whether individual traits (e.g. principal licensure exam scores) and/or environmental characteristics (e.g. becoming a principal in the same school level once served in as an

assistant principal) exert an influence on principal effectiveness. Below, we briefly detail each of the focal variables employed in value-added models.

Individual characteristics: The individual principal characteristics included in value-added models can be broken into three categories: (1) demographic traits; (2) human capital indicators; and (3) principal preparation measures.

- Principal gender: A dichotomous indicator for female principals.

- Principal ethnicity: A dichotomous indicator for minority (non-white) principals.

- College Selectivity: Dichotomous indicators for the selectivity (Barron's ranking) of the undergraduate institution early-career principals attended. Following the coding scheme of Clotfelter and colleagues, those who graduated from *not competitive* or *less competitive* institutions of higher education serve as the reference group, while we include a covariate for graduates of *competitive* institutions and a single indicator for graduates of *very competitive*, *highly competitive*, or *most competitive* institutions (Clotfelter, Ladd, & Vigdor, 2007, 2010).

- National Board Certification: A dichotomous indicator for whether an individual was a Nationally Board Certified teacher prior to assuming the principalship. Research from North Carolina indicates that National Board Certification status is a signal of teacher quality (Clotfelter, Ladd, & Vigdor, 2007, 2010; Goldhaber & Anthony, 2007). Results from this work will suggest whether this may also be a signal of principal effectiveness.

- Principal licensure exam scores: A standardized variable for a principal's performance on state required principal licensure exams. Research from North Carolina indicates that teacher licensure exam scores or a composite of all prior exam scores significantly predicts teacher quality (Clotfelter, Ladd, & Vigdor, 2007, 2010; Goldhaber, 2007; Henry, Thompson, Bastian, Fortner, Kershaw, Purtell, & Zulli, 2010). Results from this work will suggest whether principal exams may similarly signal effectiveness.[7]

- North Carolina Principal Fellow: A dichotomous indicator for whether an individual is a North Carolina Principal Fellow—an example of a principal preparation program with components expected to improve principal

[7] In addition to this standardized measure of principal exam scores, we also included specifications with a dichotomous indicator for principals with exam scores in the top quintile of the principal exam score distribution.

performance (Milstein & Krueger, 1997). The North Carolina Principal Fellows is a competitive, merit-based scholarship program that provides funding for individuals to attend school on a full-time basis and earn a Master's of School Administration from an in-state public university in two years—the first year of the program is full-time study while the second year is a full-time internship. Scholarship recipients are required to maintain employment as an assistant principal or principal in North Carolina public schools for at least four years or repay part (all) of the scholarship monies.

- Principal preparation degree type: A set of dichotomous indicators comparing the reference group—individuals whose principal preparation was a master's degree from an in-state public university—to those who earned the following principal preparation degrees: (1) a master's degree from an in-state private university; (2) a master's degree from an out-of-state university; (3) a doctorate from an in-state public university; (4) a doctorate from an in-state private university; or (5) a doctorate from an out-of-state university. Research from North Carolina indicates that degree origin matters; in particular, teachers traditionally prepared out-of-state are less effective than their in-state prepared peers (Henry, Bastian, Fortner, Kershaw, Purtell, Thompson, & Zulli, 2013). Results from this work will suggest whether degree origin may also influence principal effectiveness.

Environment characteristics: The focal variables considered in this category examine whether a principal's familiarity with or exposure to (1) a specific school (2) a specific school level or (3) a high value-added school is significantly associated with student achievement gains once principals assume leadership.

- Teacher-school match: A dichotomous indicator for whether an individual became a principal in the same school in which he/she once taught. Becoming a principal in the same school may signal a familiarity with the environment which benefits effectiveness; conversely, familiarity with the teachers as a peer may make a principal less effective when placed in a supervisory role. Research from New York City indicates that this type of match is detrimental to student achievement but is positively related to other school outcomes, such as suspension rates (Clark, Martorell, & Rockoff, 2009).

- Assistant principal-school match: A dichotomous indicator for whether or not an individual became a principal in the same school in which he/she was once an assistant principal. Becoming a principal in the same school may signal a familiarity with the environment which benefits effectiveness, and in fact, research from New York City indicates that this type of match improves student achievement in schools run by first or second year principals (Clark, Martorell, & Rockoff, 2009).

- Assistant principal-school level match: A dichotomous indicator for whether or not an individual becomes a principal at the same school level (elementary, middle, high) in which he/she was once an assistant principal. This match may indicate that a principal is more familiar with the requirements of managing students, teachers, and school operations at that level.

- Quality of the assistant principal learning environment: Research in the field of teacher quality shows that the school environment in which student teachers work and learn significantly affects their performance as novice teachers. Those individuals who student taught in "model environments"—schools that were easier-to-staff and with greater teacher collaboration—were more effective beginning teachers (Ronfeldt, 2012). Conceptually, being an assistant principal is analogous to student teaching—an apprenticeship to observe and practice the skills required of principals (vicarious and mastery experiences). Therefore, we quantify a high-quality learning environment using a measure of school value-added and test whether principals who were assistant principals (apprenticed) in a high-quality learning environment are associated with larger student test score gains than peers who were assistant principals in lower-performing schools. Specifically, we examine two measures for this analysis—a standardized value of the assistant principalship school's value-added and a dichotomous indicator for an assistant principalship school in the top quintile of school value-added. (See Appendix B for a description of how we estimated this measure of the assistant principal learning environment).

Covariates: To isolate the associations between our focal principal variables and student achievement gains, we include a rich set of student, classroom/teacher, and school covariates in our value-added models—see Table 3.1 for a complete list of available covariates. While the student and school variables are outside a principal's control, and therefore, should be included in value-added models, it is uncertain to what degree principals have discretion over the composition of their school's teacher workforce. If principals have the autonomy to make human capital decisions (hiring, firing, teacher assignment) then controlling for teacher characteristics—e.g. preparation, experience, out-of-field status—moderates a key pathway by which principals influence student outcomes and therefore, attenuates the effect estimate. Conversely, if principals

lack the authority to shape their teaching staffs, whether due to school district policies, labor market factors, or perhaps most pertinent for this analysis, principals' novice status, then controlling for teacher characteristics may be warranted. Therefore, we provide results from value-added specifications where we both exclude (preferred specification) and include (specification checks) teacher and classroom characteristics.

ANALYSIS PLAN

Although significant methodological and conceptual difficulties exist, education researchers have initiated a nascent research agenda to estimate principal effectiveness and identify its determinants using an education production function. In the present study we contribute to this developing body of work by examining the associations between individual principal characteristics or those of the environments in which they previously worked and student achievement. Specifically, we employ a value-added model with a rich set of covariates and cluster adjusted standard errors at the school-year level to estimate the relationship between our focal principal variables and adjusted-average student test score gains. Main models include all principal-school year combinations from our sample of four cohorts of first-time principals (first through fifth year principals). To determine whether associations differ according to principal experience, specification checks are limited to principals in their (1) first and second year and (2) third through fifth year.[8] The basic equation used for these models is as follows:

$$A_{ijst} = \alpha A_{it-n} + \beta Prin\_char_{st} + \gamma X_{ijst} + \delta C_{jst} + \theta S_{st} + \varepsilon_{ijst} \quad (1)$$

where $A_{ijst}$is the test score for student *i* in classroom *j* in school *s* at time *t*;

[8] All models include single year indicators of principal experience. In main models first year principals are the reference group and we include indicators for second through fifth year principals. In models limited to first and second year principals we include an indicator for second year principals. Finally, in models limited to third through fifth year principals, third year principals are the reference group and we include indicators for fourth and fifth year principals.

A_{it-n} represents the prior test score(s) for student *i*;

$Prin\_char_{st}$ is a focal principal (individual or environmental) characteristic of interest;

β estimates the association between the principal characteristic and students' adjusted-average test score growth;

X_{ijst} represents a set of time-invariant and varying individual student characteristics;

C_{jst} represents a set of classroom and teacher characteristics (excluded from main models, included in specification checks);

S_{st} represents a set of school characteristics;

and ε_{ijst} is a disturbance term representing all unexplained variation in student achievement.

Overall, to consider these focal estimates causal requires an assumption that principal assignment to schools is not confounded by variables that (1) affect both principal assignment and student test scores and (2) are excluded as covariates from the model. One might expect that early-career principals with more human capital are positively sorted into schools—for instance, that principals scoring higher on licensure exams assume leadership at higher-performing schools. To adjust for these potential confounding variables we include a rich-set of covariates. Since it is difficult to adjust for all potential confounders and to disentangle the effect of a principal (or a principal characteristic) from that of the school, prior researchers have also used school and/or principal fixed effects to reduce internal validity threats. While these fixed effects may remove confounding school variables, identification is then based on principal/school transitions, which may sharply reduce the estimation sample, often to a small network of principals, and raise the potential for bias from non-probability sample selection (Imai, Stuart, & King, 2008). Therefore, in the present study, which we consider to be

foundational and the basis for hypothesis generation, we eschew school or principal-school fixed effects.

In response to these methodological concerns we acknowledge that the assumptions needed to consider our estimates causal may be tenuous, and thus, we refer to our estimates as *associations* between individual principal characteristics or characteristics of their prior work environments and student achievement gains. In this nascent research field we contend that these associations provide the basis for empirical hypotheses upon which future, more methodologically rigorous research can build. To best isolate the relationships between principal characteristics and student achievement gains in the meantime, we specify models excluding/including teacher and classroom covariates and run separate models for first and second year principals and those principals in their third through fifth year. In addition, due to the "homegrown" nature of principals—many were teachers and/or assistant principals within the same school districts in which they become principals—and the possibility that labor market factors or district policies influence principal quality, we perform robustness checks with school district fixed effects. Results from these fixed effects specifications should be interpreted cautiously, as sample sizes experiencing within-unit (district) variation for a focal covariate may be small within many school districts.

FINDINGS

What are the characteristics of first-time principals and the schools that hire them?

To better understand the characteristics of those who become principals in North Carolina public schools we identified four cohorts of first-time principals (from 2006-07 through 2009-10) and used administrative data to detail their demographics, prior work

experiences, principal preparation, and the types of schools in which they first assume leadership. We display these data in Tables 3.2 and 3.3 and discuss relevant findings in the sub-sections below.

Demographics: As displayed in the top panel of Table 3.2, our full sample is comprised of 981 first-time principals whose average age at first assuming school leadership is 41 years and whose standardized principal licensure exam scores are one-fifth of a standard deviation above the mean for all test-takers. Overall, nearly 60 percent of first-time principals are female and seventy-three percent are white.[9] While average age at first principalship is fairly comparable across gender and ethnicity, there are sizable differences in principal licensure exam scores—on average, female and white principals outscore their male and minority peers, respectively, by more than one-third of a standard deviation.

Career Trajectories: As shown in the bottom left panel of Table 3.2, 880 (89.70%) of the 981 first-time principals in our sample were previously teachers in North Carolina public schools. On average, these individuals worked as teachers in the state's public schools for nearly 8.5 years and as teachers overall for nearly 11 years. Sixty-five percent of these individuals served as teachers at the same school level (elementary, middle, or high school) in which they became first-time principals. Approximately 11 percent of these individuals held National Board Certification;[10] the most common teaching license was in elementary grades, followed by secondary grades English and

[9] Overall, ethnicity data is available for 978 of the 981 first-time principals in our sample. Of the 265 minority principals 245 are black, 13 are Native American, 2 are Hispanic, 2 are Asian, and 3 are classified as other.

[10] As a basis of comparison for the National Board Certification percentage in our principal sample, in the 2004-05 school year (when many of the principals in our sample had already transitioned into the assistant principalship) only 7.08% of the teachers in North Carolina public schools were Nationally Board Certified.

social studies. Regarding career trajectories as assistant principals, the bottom right panel of Table 3.2 shows that 918 (93.58%) of the 981 first-time principals in our sample were previously assistant principals in North Carolina public schools. On average, these individuals worked as assistant principals in the state's public schools for 4.15 years (with a range from 1 year to 18 years). Nearly three-fourths of these individuals served as assistant principals at the same school level in which they became first-time principals.

"Homegrown" Principals: Perhaps the most striking career trajectory data concerns the propensity of school districts to promote individuals to the principalship from within the district. Of the 880 first-time principals in our sample who taught in North Carolina public schools, nearly 75 percent taught in the same school district in which they became principal and more than 11 percent taught in the same school—on average, individuals served as teachers in these environments for 7.57 and 6.06 years, respectively. Almost 90 percent of the 918 individuals who served as assistant principals did so in the same school district in which they became a first-time principal and nearly a quarter transitioned to the principalship within the same school. The average time spent as assistant principals in these environments was 3.86 and 3.26 years, respectively. Finally, 586 individuals—nearly 60 percent of our sample—became a first-time principal in the same school district in which they had both served as a teacher and assistant principal; on average, the combined years of service in these districts prior to assuming the principalship was 11.45 years.

Principal Preparation: Concerning principal preparation, the bottom right panel of Table 3.2 displays the last degree individuals earned prior to assuming the principalship. Overall, (1) 11 percent of our sample were North Carolina Principal

Fellows; (2) two-thirds of the first-time principals graduated with a master's degree from an in-state public institution; (3) the last degree for nearly 89 percent of our sample was a master's; and (4) approximately 14 percent of individuals earned degrees out-of-state. On average, there were just over five years between the last degree and assuming school leadership, indicating that a typical individual from our sample completed principal preparation slightly before or concurrent with becoming an assistant principal.

School Descriptives: Beyond investigating demographics and career trajectories for our sample of first-time principals, we also examined characteristics—academic performance, student composition, teacher workforce credentials—of the schools in which first-time principals assume leadership.[11] Table 3.3 presents school characteristics for four comparisons of interest: (1) first-time principals versus more experienced principals; (2) white versus minority first-time principals; (3) first-time principals with top quintile versus bottom quintile licensure exam scores; and (4) first-time principals who were assistant principals in high value-added schools (top quintile) versus those from low value-added schools (bottom quintile).

Overall, each of these comparisons indicates that principals sorted across schools according to observed characteristics. For example, in comparison to more experienced peers, first-time principals assume leadership at schools with (1) lower performance composite values (2) higher concentrations of students qualifying for subsidized school lunches (high-poverty) and (3) less well-credentialed teachers. Table 3.3 also shows that many of these differences extend *within* our sample of first-time principals as (1) white and top-scoring principals take control of higher-performing schools with better-

[11] Because principals can influence many of the characteristics presented in Table 3.3 (e.g. percentage of AYP goals met or teacher retention rates), we display school characteristics from the year prior to a principal from our sample assuming leadership.

credentialed teacher workforces and fewer minority or high-poverty students and (2) those who apprenticed (assistant principals) in high value-added schools assume leadership in higher-performing schools with fewer high-poverty students. How much these differences in first-time principal assignments are due to the preferences of individual principals, the hiring practices of school districts, or a combination of the two is unknown. It is clear, however, that these differences indicate sorting on observed characteristics which may also suggest sorting on unobserved variables and underscore our use of associational language to describe our value-added estimates.

What Principal Characteristics Are Associated with Student Achievement Gains?

Individual Characteristics: Overall, the individual principal characteristics included in value-added models can be broken into three categories: (1) demographic traits; (2) human capital indicators; and (3) principal preparation measures. Regarding demographics, Table 3.4—which presents findings for our full sample of principals from rich covariate OLS (odd numbered columns) and school district fixed effect (even numbered columns) models—indicates that female principals are no more or less effective than male principals and minority principals are no more or less effective than white principals. Additional models interacting a principal's minority status with a student's minority status, to test whether minority students make larger achievement gains under the leadership of a minority principal, return a pattern of negative coefficients in elementary and middle grades reading and positive associations in high school but very few statistically significant results.[12] Concerning human capital indicators, results from school district fixed effect models in middle school and rich covariate adjustment models

[12] There were positive interaction coefficients for high school models limited to first and second year principals with school district fixed effects. All these interaction results are not tabled in this manuscript but are available upon request from the author.

in high school return negative and significant associations between undergraduate campus selectivity and adjusted-average student achievement gains. However, principal licensure exam scores (both a standardized and dichotomous measure), an indicator of human capital more proximate and related to becoming a principal, are positively and significantly associated with student achievement gains in elementary and middle grades mathematics. Finally, results regarding formal principal preparation degrees show that: (1) student achievement gains are significantly smaller in elementary schools run by North Carolina Principal Fellows—this is particularly noteworthy because over 75 percent of the Principal Fellows in our sample lead schools with elementary tested-grades—and (2) individuals earning Masters degrees from in-state public institutions (reference category) are generally associated with achievement gains no different from or significantly larger than (seven comparisons) those holding a doctorate prior to assuming school leadership.

Environment Characteristics: The focal variables considered in this category examine whether a principal's familiarity with or exposure to (1) a specific school (2) a specific school level or (3) a high value-added school is significantly associated with student achievement gains once principals assume leadership. As before, results in odd numbered columns in Table 3.5 are from rich covariate adjustment models excluding teacher and classroom covariates; coefficients in even numbered columns include school district fixed effects. Here, Table 3.5 shows that taking control of a school in which a principal once taught is negatively and significantly associated with student achievement gains in elementary grades mathematics and within districts in middle grades reading. This finding corroborates prior work by Clark and colleagues in New York City and may

indicate the challenges of transitioning from a subordinate to a leadership position within a school. Conversely, there is some indication that in middle grades transitioning from an assistant principal to the principal within the same school is associated with positive student outcomes, suggesting that familiarity with leadership in that environment may facilitate principal success. Even stronger across all secondary grades models—middle grades and high school—is the relationship between working as an assistant principal in a middle/high school and then assuming the principalship at the same school level. No such relationship exists in elementary schools, suggesting that there may be unique aspects of secondary grades environments that are learned through experience in a leadership position at that level. Finally, the strongest and most consistent value-added result concerns the relationship between a measure of the apprenticeship (assistant principal) learning environment quality and student achievement gains. Across all grade levels and subjects, principals who previously worked as an assistant principal in a high value-added school—a measure for an environment in which to best learn the knowledge and practices of effective school leadership—were associated with positive and significant student achievement gains. The strength of this finding and its implications for policy suggest that it may be a plausible, evidence-based hypothesis to test through more rigorous econometric methods or direct manipulation in a field experiment.

Specification Checks: In Appendix C we present results from two types of specification checks: (1) value-added models that include teacher and classroom covariates (Appendix Tables C1 and C2) and (2) value-added models (excluding teacher/classroom variables) focused on principals in their first and second year only or third, fourth, and fifth year only (Appendix Tables C3-C6). This second round of checks

examines whether the relationship between principal characteristics and student achievement gains varies with principal experience.[13]

Overall, results in Appendix Tables C1 and C2 display few differences from the preferred specifications shown in Tables 3.4 and 3.5. For our sample of first-time principals, controlling for teacher and classroom characteristics has little influence on the associations of interest. However, for the associations between student achievement gains and individual or environmental characteristics a few notable differences do exist based on principal experience. Regarding individual principal characteristics, the magnitude of some relationships dissipates as principals acquire more experience—NBC status or licensure exam scores in elementary grades—while for others significant associations develop over time—North Carolina Principal Fellows in elementary mathematics (negative) or licensure exam scores in middle grades (positive). Concerning characteristics of the work environment, the relationship between student outcomes and ever serving as an assistant principal in the same school or school level strengthens for middle school principals over time; this school level association dissipates for high school principals. Most importantly, Appendix Tables C5 and C6 show that while significant associations persist for the quality of the apprenticeship learning environment, its relationship with student achievement is much stronger in principals' first two years.

DISCUSSION

Nascent empirical research indicates that principal effectiveness significantly influences student achievement, teacher retention and development, and school working

[13] While the acquisition of on-the-job learning is one reason why results for principal characteristics may differ between the 1-2 year versus 3-5 year models, differential principal attrition—e.g. principals with the highest or lowest licensure exam scores exiting the principalship—or differential assignment patterns—e.g. school districts assigning the principals with the highest licensure exam scores to the lowest-performing schools—may also contribute to differences across models.

conditions. (Branch, Hanushek, & Rivkin, 2012; Ladd, 2011; Loeb, Kalogrides, & Beteille, 2012). This suggests that principal quality may represent an important policy lever by which to improve outcomes of interest, yet researchers, practitioners, and policymakers do not thoroughly understand the determinants of principal quality and/or why some principals succeed and others do not. Given the prevalence of inexperienced principals in the workforce (see Figure 3.1) and the drops in student achievement associated with a school's transition to a first-time principal (see Figure 3.2), further research is needed to better understand the performance of early-career school leaders. In this study we contributed to this developing research agenda by examining first-time principals, their pathways to the principalship, and the extent to which individual principal characteristics or those of their prior work environments were significantly associated with student achievement gains.

Overall, we found that a large majority of first-time principals were "homegrown"—assuming school leadership in the same districts in which they once taught and/or served as an assistant principal. This finding corroborates prior work by Brewer, which indicated that districts in New York State promoted a majority of their assistant principals and principals from positions within the district (Brewer, 1996). Districts seem to prefer having familiarity with those elevated to principal positions; however, this suggests that districts select candidates from a limited pool. Further descriptive evidence showed that based on observed characteristics, such as minority status or measured human capital, first-time principals sort into schools with different levels of academic performance, student body composition, and/or teacher workforce credentials. Whether this sorting is the result of first-time principals with more desirable

characteristics selecting into more attractive environments, the assignment patterns of districts, or a combination of both is unknown. Based on our value-added findings, which show that principal characteristics, such as licensure exam scores or the quality of the apprenticeship learning environment, are significantly associated with student achievement gains, this indicates that lower-performing schools, on average, lack access to higher-quality school leaders.

Concerning our value-added analyses, we found that several individual principal characteristics were significantly associated with student achievement gains—notably principal licensure exam scores (positive), indicating that higher levels of human capital may benefit principal performance, and North Carolina Principal Fellows (negative). Given that the Principal Fellows program is a highly recommended training model, this suggests that we currently do not know enough about effective principal preparation. Our strongest value-added results concerned characteristics of early-career principals' prior work environments. In elementary grades (math), early-career principals who assume school leadership at the same school in which they once taught are associated with negative student achievement gains. This result is consistent with work by Clark and colleagues in New York City and indicates that it may be difficult for early-career principals to exert authority over teachers who were previously co-workers (Clark, Martorell, & Rockoff, 2009). In secondary grades congruence between the level of the assistant principal and principal schools was significantly associated with achievement gains—no such relationship exists for this congruence in elementary schools. This may suggest that there are unique aspects of secondary grades environments—larger student bodies, subject-level departments, higher concentrations of alternatively prepared

teachers—that (1) provide assistant principals with greater opportunities to exercise leadership and (2) are best learned through specific exposure to and practice within this schooling level. Across all grade levels, the value-added effectiveness of the school in which early-career principals served their apprenticeship (assistant principal) strongly predicted principal effectiveness. Coupled with the finding that, on average, five years pass between the formal principal preparation degree and assuming the principalship, these associations suggest that the effectiveness of early-career principals may hinge on their apprenticeship environment quality.

From a research perspective we acknowledge that the assumptions needed to treat these findings as causal may be tenuous—for instance, that (1) the assignment of a high-scoring principal to a school is uncorrelated with omitted school characteristics (such as prior school performance) and (2) these omitted school characteristics are uncorrelated with student-level achievement. Our estimates do provide an empirical foundation for hypothesis generation and future research initiatives. Two specific examples illustrate this point. First, North Carolina annually invests millions of dollars into its Principal Fellows program, yet results from this research indicate that Principal Fellows are associated with significantly reduced student achievement outcomes in elementary schools—where Principal Fellows are highly concentrated. This suggests the need for a more rigorous evaluation to estimate the effects of Principal Fellows; results from this evaluation may signal the need to develop additional principal preparation models and test their effects on outcomes of interest. Second, this study suggests a place to look for improving principal performance—the quality of the apprenticeship learning environment. To better understand this finding and to determine whether it is due to on-

the-job learning as an assistant principal (mastery and vicarious experiences in high value-added schools) or early-career principals sorting into better schools, future research should: (1) examine the process by which individuals are assigned to their apprenticeship and principalship schools; (2) test additional measures of a high-quality learning environment for an assistant principal (e.g. teacher survey responses regarding school culture and instructional leadership); and (3) employ more sophisticated methods, such as differences-in-differences with school fixed effects, to isolate the apprenticeship learning environment effect. Following Bandura's model of self-efficacy formation, it would be beneficial to understand how apprenticeships (1) provide opportunities for mastery experiences, in which the assistant principal directly engages in practices that improve teacher and student performance, and vicarious experiences, in which sitting principals model effective practices and (2) enhance the attribution of student success to practices that the principal and assistant principals enact (Bandura, 1977; Hoy & Spero, 2005; Mulholland & Wallace, 2001). To disentangle these effects of environment from those of human capital, it will likely be necessary to design random assignment studies in which the assignment of assistant principals to school environments is intentionally manipulated.

From a policy perspective, we believe that conceptualizing the determinants of principal effectiveness in these two broad categories—individual characteristics and those of prior work environments—provides a clear focus for future policy. If individual characteristics, such as human capital indicators (e.g. principal licensure exam scores) or preparation type matter, this suggests policy levers targeted at recruiting and selecting individuals with those desired qualities. If previous work experiences matter, as the

findings from this research, particularly the apprenticeship learning environment, suggest, this indicates that effective school leadership may be learned through exposure and that districts' patterns of assigning assistant principals and principals to schools should be carefully considered. As the principal effectiveness research agenda develops, these implications for policy should become clearer and facilitate policy action to improve early-career principal performance.

REFERENCES

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215.

Beteille, T., Kalogrides, D., & Loeb, S. (2012). Stepping stones: Principal career paths and school outcomes. *Social Science Research, 41*(4), 904-919.

Boyd, D., Grossman, P., Ing, M., Lankford, H., & Loeb, S. (2011). The influence of school administrators on teacher retention decisions. *American Educational Research Journal, 48*(2), 303-333.

Branch, G., Hanushek, E., & Rivkin, S. (2012). Estimating the effect of leaders on public sector productivity: The case of school principals. *NBER Working Paper 17803.*

Brewer, D. (1993). Principals and student outcomes: Evidence from U.S. high schools. *Economics of Education Review, 12*(4), 281-292.

Brewer, D. (1996). Career paths and quit decisions: Evidence from teaching. *Journal of Labor Economics, 14*(2), 313-339.

Burkhauser, S., Gates, S., Hamilton, L., & Ikemoto, G. (2012). First-year principals in urban school districts: How actions and working conditions relate to outcomes. *RAND Education.*

Chiang, H., Lipscomb, S., & Gill, B. (2012). Is school value-added indicative of principal quality. *Mathematica Policy Research.* Available from: http://mathematica-mpr.com/publications/pdfs/education/value-added_principal_quality.pdf.

Clark, D., Martorell, P., & Rockoff, J. (2009). School principals and school performance. *Calder Institute Working Paper 38.*

Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal Analysis with student fixed effects. *Economics of Education Review, 26*(6), 673-682.

Clotfelter, C., Ladd, H., & Vigdor, J. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources, 45*(3), 655-681.

Coelli, M., & Green, D. (2012). Leadership effects: School principals and student outcomes. *Economics of Education Review, 31*(1), 92-109.

Corcoran, S., Schwartz, A.E., & Weinstein, M. (2012). Training your own: The impact of New York City's Aspiring Principals Program on student achievement. *Educational Evaluation and Policy Analysis, 34*(2), 232-253.

Dhuey, E., & Smith, J. (2012). How school principals influence student learning. Available from: http://homes.chass.utoronto.ca/~edhuey/datastore/files/docs/Dhuey_Smith_Nov_12_2012.pdf.

Eberts, R., & Stone, J. (1988). Student achievement in public schools: Do principals make a difference? *Economics of Education Review, 7*(3), 291-299.

Goddard, R.D., Hoy, W.K., & Woolfolk Hoy, A. (2000). Collective teacher efficacy: Its meaning, measure and effect on student achievement. *American Educational Research Journal, 37*(2), 479-507.

Goldhaber, D. (2007). Everyone's doing it, but what does teacher testing tell us about teacher effectiveness? *Journal of Human Resources, 47*(1), 765-794.

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.

Grissom, J., & Loeb, S. (2011). Triangulating principal effectiveness: How perspectives of parents, teachers, and assistant principals identify the central importance of managerial skills. *American Educational Research Journal, 48*(5), 1091-1123.

Grissom, J., Kalogrides, D., & Loeb, S. (2012). Using student test scores to measure principal performance. *NBER Working Paper 18568.*

Handa, S., Thompson, C.L., Marcus, J.V., & Smith, A.A. (2010). The effects of UNC Master of School Administration programs on NC public schools, 2004-2008. *Carolina Institute for Public Policy.* Available from: http://publicpolicy.unc.edu/research/publications-presentations-and-reports.

Henry, G.T., Thompson, C.L., Bastian, K.C., Fortner, C.K., Kershaw, D.C., Purtell, K.M., & Zulli, R. A. (2010). Portal report: Teacher preparation and student test scores in North Carolina. *Carolina Institute for Public Policy.* Available from: http://publicpolicy.unc.edu/research/publications-presentations-and-reports.

Henry, G.T., Bastian, K.C., Fortner, C.K., Kershaw, D.C., Purtell, K.M., Thompson, C.L., & Zulli, R.A. (2013). Teacher preparation policies and their effects on student achievement. In press, *Education Finance and Policy*.

Horng, E., Klasik, D., & Loeb, S. (2010). Principal's time use and school effectiveness. *American Journal of Education, 116*(4), 491-523.

Hoy, A.W., & Spero, R.B. (2005). Changes in teacher efficacy during the early years of teaching: A comparison of four measures. *Teaching and Teacher Education, 21*(6), 343-356.

Imai, K., King, G., & Stuart, E. (2008). Misunderstandings between experimentalists and observationalists about causal inference. *Journal of the Royal Statistical Society: Series A, 171*(2), 481-502.

Ladd, H. (2011). Teachers' perceptions of their working conditions: How predictive of planned and actual teacher movement? *Educational Evaluation and Policy Analysis, 33*(2), 235-261.

Loeb, S., Kalogrides, D., & Horng, E.L. (2010). Principal preferences and the uneven distribution of principals across schools. *Educational Evaluation and Policy Analysis, 32*(2), 205-229.

Loeb, S., Kalogrides, D., & Beteille, T. (2012). Effective schools: Teacher hiring, assignment, development, and retention. *Education Finance and Policy, 7*(3), 269-304.

Martorell, F., Heaton, P., Gates, S., & Hamilton, L. (2010). Preliminary findings from the New Leaders for New Schools evaluation. *RAND Corporation.* Available from: http://www.rand.org/content/dam/rand/pubs/working_papers/2010/RAND_WR739.pdf.

Mulholland, J., & Wallace, J. (2001). Teacher induction and elementary science teaching: Enhancing self-efficacy. *Teaching and Teacher Education, 17*(2), 243-261.

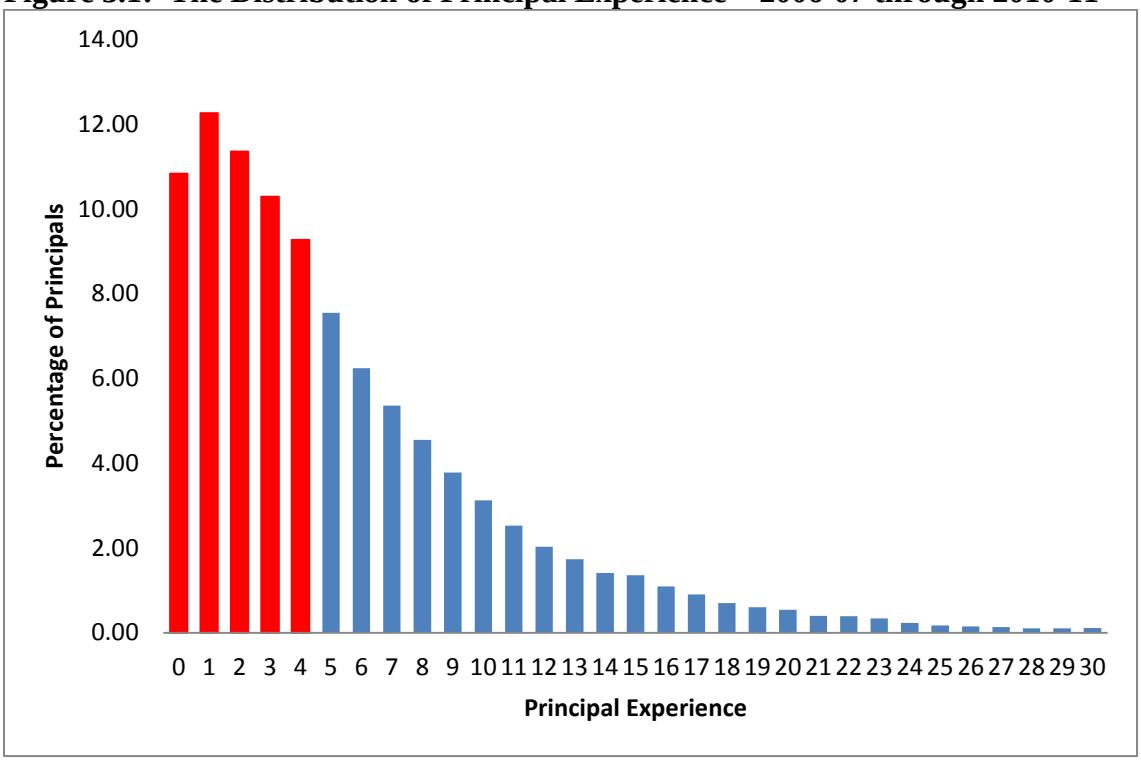
Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 533-578.

Robinson, V., Lloyd, C., & Rowe, K. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly, 44*(5), 635-674.

Ronfeldt, M. (2012). Where should student teachers learn to teach? Effects of field placement school characteristics on teacher retention and effectiveness. *Educational Evaluation and Policy Analysis, 34*(1), 3-26.

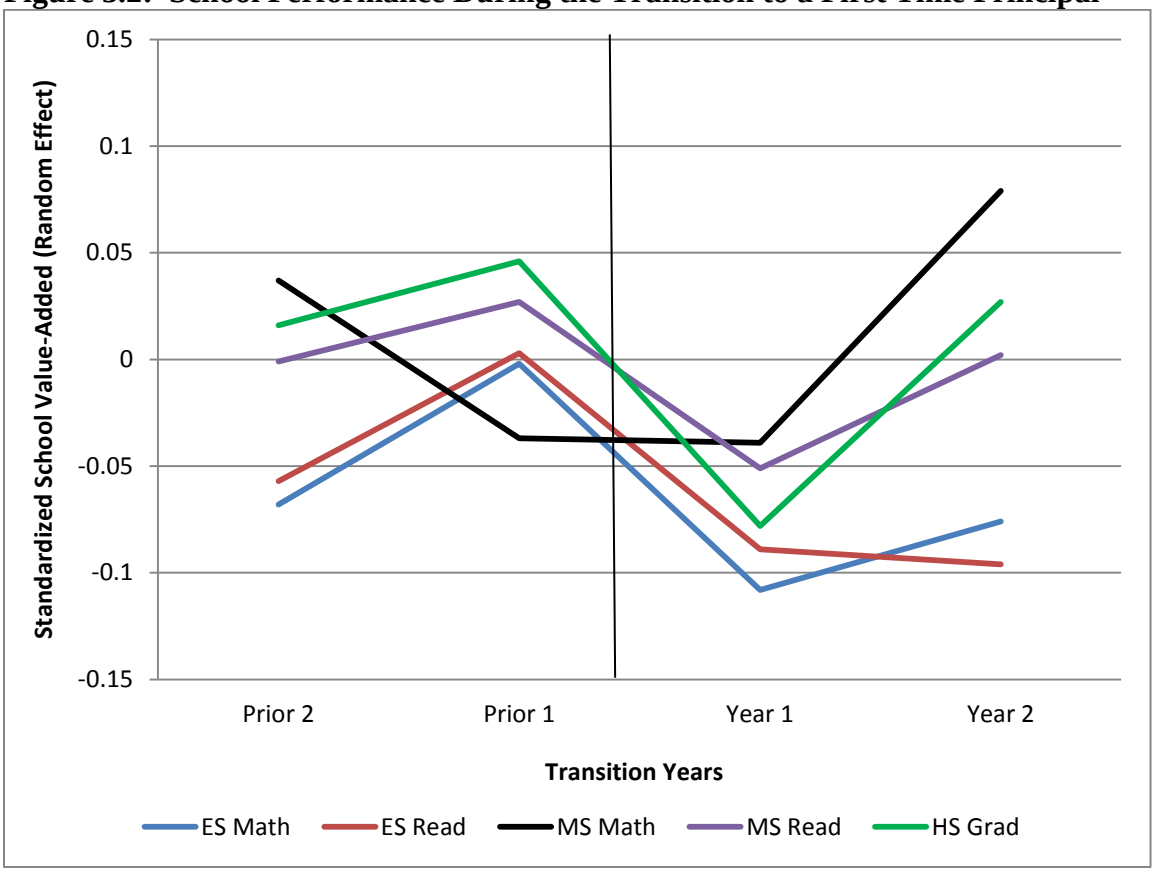
Vanderhaar, J., Munoz, M. & Rodosky, R. (2006). Leadership as accountability for learning: The effects of school poverty, teacher experience, previous achievement, and principal preparation programs on student achievement. *Journal of Personnel Evaluation in Education, 19*(1-2), 17-33.

Figure 3.1: The Distribution of Principal Experience—2006-07 through 2010-11



Note: This figure displays the levels of principal experience for all North Carolina public school principals in the 2006-07 through 2010-11 school years.

Figure 3.2: School Performance During the Transition to a First Time Principal



Note: This figure displays standardized school effectiveness values (school value-added) for the two years prior to and two years immediately after transition to a first-time principal. The vertical line in the center of the graph marks the transition to a first-time principal.

Table 3.1: Covariates Available for Value-Added Models

Student Covariates	Classroom/Teacher Covariates	School/Principal Covariates
1) Prior student test scores 2) Peer ability 3) Days absent 4) Structural mobility 5) Within year mobility 6) Between year mobility 7) Underage for grade 8) Overage for grade 9) Giftedness 10) Disability 11) Free or reduced-price lunch 12) Ethnicity 13) Gender 14) Currently limited English proficient 15) Was limited English proficient 16) Year fixed effects 17) Course indicators (HS only)	1) Class size 2) Heterogeneity of prior student performance 3) Advanced curriculum (MS and HS only) 4) Remedial curriculum (MS and HS only) 5) Out-of-field teaching 6) Teacher preparation type 7) Teacher experience	1) School size 2) School size squared 3) Total per-pupil expenditures 4) Average teacher supplement 5) Short-term suspension rate 6) Violent acts rate 7) Free and reduced-price lunch percentage 8) Race/ethnicity percentages 9) Focal principal characteristics a. Individual trait b. Environmental trait 10) Principal experience

Note: All value-added specifications include the covariates included in columns 1 and 3 (student and school/principal). Classroom and teacher covariates are excluded from the main results presented in Tables 3.4 and 3.5 but included in specification checks.

Table 3.2: Demographic Information and Career Trajectory

Demographic Information					
	Full Sample	**Female Principals**	**Male Principals**	**White Principals**	**Minority Principals**
Count	981	586 (59.73%)	395 (40.27%)	713 (72.90%)	265 (27.10%)
Average Age at 1st Principalship	41.37	42.63	39.51	41.58	40.73
Std. Principal Exam Scores	0.192	0.344	-0.035	0.327	-0.032

Teacher Career Trajectory		**Assistant Principal Career Trajectory and Principal Preparation**	
Previously a teacher in NCPS	880	Previously an assistant principal (AP) in NCPS	918
Average years as a teacher in NCPS	8.40	Average years as an AP in NCPS	4.15
Average years as a teacher overall	10.98		
Teacher in the same LEA	643/880 (73.07%)	Assistant principal in the same LEA	817/918 (88.99%)
Average years as a teacher in the same LEA	7.57	Average years as an AP in the same LEA	3.86
Teacher in the same school	100/880 (11.36%)	Assistant principal in the same school	216/918 (23.53%)
Average years as a teacher in the same school	6.06	Average years as an AP in the same school	3.26
Teacher in the same school type	571/880 (64.89%)	Assistant principal in the same school type	672/918 (73.20%)
Average years as a teacher in the same school type	7.13	Average years as an AP in the same school type	3.54
National Board Certification	97/880 (11.02%)	NC Principal Fellow	108/981 (11.01%)
Teaching Licenses Held		Average years from last degree to principalship	5.12
Elementary	41.73%	Last degree prior to beginning principalship	
Middle or High School Math	14.93%	*NC Public University Master's*	66.83%
Middle or High School English	27.92%	*NC Private University Master's*	11.83%
Middle or High School Science	15.04%	*Out of State University Master's*	9.89%
Middle or High School Social Studies	22.34%	*NC Public University Doctorate*	6.83%
Exceptional Children	18.24%	*NC Private University Doctorate*	0.51%
PE and Health Education	9.23%	*Out of State University Doctorate*	4.08%

Note: The top panel of this table displays basic demographic and test score information for our full sample of first-time principals (981 in total) and breakdowns according to gender and ethnicity. The bottom panels of this table display career trajectory (teacher and assistant principal) and principal preparation descriptives. When appropriate, cells display both counts and percentages. For teacher descriptives, percentages are in reference to the total number of teachers (880). For assistant principal descriptives, percentages are in reference to the total number of assistant principals (918).

Table 3.3: School Characteristics Inherited by First-Time Principal Sample

	Principal Experience Comparisons			White vs. Minority Principals		Principal Licensure Exam Scores		Value-Added at the Last Assistant Principal School	
School Characteristics	**Sample of 1st Time Principals**	**Principals with 5-10 Years Exp.**	**Principals with More than 10 Years Exp.**	**White 1st Time Principals**	**Minority 1st Time Principals**	**Top-Scoring 1st Time Principals**	**Bottom-Scoring 1st Time Principals**	**1st Time Principals from High VA Schools**	**1st Time Principals from Low VA Schools**
Performance Composite (# of tests passed/# tests taken)	58.14	60.01	62.35	61.35	49.60	60.78	55.09	61.67	55.30
Percentage of AYP Goals Met	88.15	88.94	88.57	90.01	83.19	89.08	84.91	90.62	85.65
NC ABC Growth Status									
No Growth	29.94	25.90	25.08	28.61	33.88	29.34	31.21	20.25	42.86
Expected Growth	40.11	39.97	40.20	39.73	41.22	37.72	39.31	39.24	33.08
High Growth	29.94	34.12	34.72	31.66	24.90	32.93	29.48	40.51	24.06
Short-term Suspension Rate (Per 100 students)	25.11	23.24	21.29	20.76	36.72	20.48	36.56	27.73	26.37
Percentage of Minority Students	47.10	44.50	46.46	38.72	69.01	43.75	52.27	48.17	49.98
Percentage of Subsidized School Lunch Students	56.23	52.53	50.33	51.77	67.93	50.67	61.88	52.61	62.56
Total Per-Pupil Expenditures	$9,263	$9,266	$9,456	$8,990	$9,998	$9,360	$9,856	$9,314	$9,301
Average Teacher Supplements	$2,837	$2,966	$3,152	$2,752	$3,060	$3,208	$2,647	$3,043	$2,592
Percentage of NBC Teachers	8.59	9.54	9.88	9.54	6.08	9.66	6.42	8.73	7.59
Percentage of Novice Teachers	24.06	22.48	21.85	22.94	26.98	24.77	24.90	24.75	24.75
Percentage of Teachers with an Advanced Degree	26.75	27.71	29.42	27.20	25.54	27.28	24.85	27.02	24.93
Percentage of Teachers Returning to the School	76.67	79.72	80.25	77.92	73.35	76.92	74.75	75.50	75.77
Total Number of Principal-School-Year Combinations	905	2,325	1,660	657	245	167	173	158	133

Note: This table displays school-level data from the year before our sample of 981 principals are appointed (descriptive data for the schools our sample inherits). In total there are 905 principal-school combinations for our sample, indicating that 76 principals assume leadership at brand new schools (no previous year school-level data). Overall, this table shows four comparisons detailed in the results section.

Table 3.4: Principal Characteristics and Student Achievement (Full Sample)

Focal Variable	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Female Principal	-0.008	-0.007	0.008	0.002	-0.001	-0.003	0.005	0.005	0.011	-0.013
	(0.007)	(0.007)	(0.005)	(0.005)	(0.007)	(0.008)	(0.004)	(0.004)	(0.014)	(0.015)
Cases	566,293		770,983		423,328		446,125		428,450	
Minority Principal	-0.010	-0.007	-0.002	0.000	**-0.029****	-0.015	0.001	-0.005	0.013	-0.007
	(0.008)	(0.008)	(0.007)	(0.006)	(0.009)	(0.010)	(0.006)	(0.006)	(0.022)	(0.022)
Cases	564,601		768,683		423,328		446,125		426,884	
Selectivity										
Competitive	-0.003	-0.011	0.003	0.011	-0.006	**-0.029***	-0.001	**-0.016***	**-0.063****	-0.024
	(0.009)	(0.008)	(0.007)	(0.007)	(0.010)	(0.012)	(0.006)	(0.007)	(0.017)	(0.020)
Highly Comp	0.002	-0.011	-0.001	**0.015***	-0.018	-0.020	-0.008	-0.009	**-0.078****	-0.027
	(0.010)	(0.010)	(0.007)	(0.007)	(0.012)	(0.014)	(0.006)	(0.008)	(0.019)	(0.021)
Cases	566,293		770,983		423,328		446,125		428,450	
NBC	0.020	0.008	**0.016***	0.005	-0.022	-0.015	-0.012	-0.012	0.002	-0.015
	(0.011)	(0.010)	(0.008)	(0.008)	(0.014)	(0.016)	(0.007)	(0.009)	(0.026)	(0.030)
Cases	566,293		770,983		423,328		446,125		428,450	
Principal Exam Score	**0.011****	0.006	0.003	0.000	**0.009***	0.001	0.002	0.004	0.002	-0.017
	(0.004)	(0.004)	(0.005)	(0.003)	(0.004)	(0.005)	(0.003)	(0.003)	(0.006)	(0.009)
Cases	538,402		731,710		393,767		414,060		408,354	
Top Quintile Exam Score	**0.016***	0.009	**0.015***	0.003	**0.020***	0.014	0.002	0.008	0.000	**-0.054****
	(0.008)	(0.008)	(0.006)	(0.006)	(0.009)	(0.011)	(0.005)	(0.006)	(0.013)	(0.019)
Cases	538,402		731,710		393,767		414,060		408,354	

Table 3.4 (Continued): Principal Characteristics and Student Achievement (Full Sample)

Focal Variable	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
NC Principal Fellow	**-0.024***	**-0.024***	**-0.022****	-0.009	0.004	-0.028	-0.000	-0.010	0.000	0.021
	(0.009)	(0.010)	(0.007)	(0.007)	(0.016)	(0.016)	(0.007)	(0.009)	(0.021)	(0.036)
Cases	566,293		770,983		423,328		446,125		428,450	
Preparation										
Private Masters	0.003	0.007	-0.005	-0.007	-0.010	0.012	-0.002	0.011	-0.004	-0.031
	(0.012)	(0.013)	(0.008)	(0.009)	(0.012)	(0.015)	(0.006)	(0.008)	(0.017)	(0.021)
Out-of-state Masters	0.002	-0.007	0.008	0.004	0.016	**0.034****	0.013	**0.016***	-0.014	0.002
	(0.011)	(0.011)	(0.008)	(0.009)	(0.012)	(0.013)	(0.007)	(0.007)	(0.024)	(0.022)
Public Doctorate	0.007	0.003	**-0.032****	-0.007	**-0.025***	-0.018	-0.011	0.010	0.046	-0.026
	(0.013)	(0.013)	(0.009)	(0.009)	(0.012)	(0.014)	(0.008)	(0.009)	(0.028)	(0.028)
Private Doctorate	-0.035	-0.003	-0.078	**-0.120****	0.026	---	**-0.055****	---	0.003	-0.061
	(0.072)	(0.038)	(0.047)	(0.035)	(0.039)		(0.016)		(0.030)	(0.040)
Out-of-State Doctorate	**-0.034***	-0.014	-0.002	-0.008	**-0.034***	-0.008	0.014	-0.003	0.017	**-0.052***
	(0.016)	(0.019)	(0.014)	(0.014)	(0.017)	(0.027)	(0.008)	(0.008)	(0.027)	(0.026)
Cases	565,246		769,921		423,328		446,125		428,450	

Note: Odd numbered columns display results from rich covariate adjustment models (OLS) without teacher/classroom covariates. Even numbered columns display results from models with school district fixed effects and no teacher/classroom covariates. * Indicates statistical significance at the p<0.05 level; ** indicates statistical significance at the p<0.01 level.

Table 3.5: Previous Work Environment Characteristics and Student Achievement (Full Sample)

	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
Focal Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Teacher-School Match	**-0.032***	**-0.047****	0.005	0.003	-0.003	0.000	-0.004	**-0.014***	0.004	-0.008
	(0.013)	(0.012)	(0.010)	(0.009)	(0.012)	(0.011)	(0.006)	(0.007)	(0.018)	(0.022)
Cases	566,293		770,983		423,328		446,125		428,450	
AP-School Match	0.005	0.009	0.006	0.002	0.014	**0.021***	**0.012****	0.007	0.008	0.022
	(0.008)	(0.008)	(0.006)	(0.006)	(0.008)	(0.010)	(0.004)	(0.006)	(0.015)	(0.015)
Cases	566,293		770,983		423,328		446,125		428,450	
AP-School Type Match	-0.010	-0.014	-0.009	-0.003	**0.019***	**0.019***	0.008	**0.011***	**0.070****	**0.045****
	(0.008)	(0.008)	(0.006)	(0.006)	(0.008)	(0.008)	(0.004)	(0.005)	(0.019)	(0.019)
Cases	566,293		770,983		423,328		446,125		428,450	
Std. Apprenticeship Learning Environment	**0.018****	**0.012****	**0.011****	0.004	**0.022****	**0.016****	**0.009****	0.005	**0.031****	**0.022****
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)	(0.002)	(0.003)	(0.008)	(0.008)
Cases	518,843		704,197		403,579		425,511		401,981	
Top Quintile Apprenticeship Learning Environment	**0.024****	0.008	**0.020****	**0.014***	**0.039****	**0.037****	**0.019****	**0.021****	**0.062****	**0.057****
	(0.009)	(0.009)	(0.006)	(0.006)	(0.009)	(0.010)	(0.005)	(0.005)	(0.016)	(0.015)
Cases	518,843		704,197		403,579		425,511		401,981	

Note: Odd numbered columns display results from rich covariate adjustment models (OLS) without teacher/classroom covariates. Even numbered columns display results from models with school district fixed effects and no teacher/classroom covariates. * Indicates statistical significance at the p<0.05 level; ** indicates statistical significance at the p<0.01 level.

CONCLUSION

While each chapter in this dissertation focused on different aspects of teacher or principal quality, all three chapters were united by a common purpose: advancing the quality of school-based human capital and subsequently, student achievement, through the provision of research evidence identifying promising policy practices. Toward this end, my dissertation both proposed specific policy recommendations, such as reducing the adverse effects of out-of-state prepared teachers by increasing the number of teachers prepared and hired in-state, and served as a foundation for hypothesis generation and future research initiatives. Given my standing as a doctoral student transitioning towards an independent researcher, I dedicate this conclusion to discussing potential research projects stemming from this dissertation.

Regarding my first dissertation chapter, evidence suggests that pre-service measures of TFA corps members' non-cognitive characteristics predict outcomes of interest. Taken to its conclusion, this indicates that teacher preparation programs and school districts may enhance teacher quality by (1) incorporating measures of non-cognitive attributes into recruitment and selection/hiring decisions and (2) cultivating these characteristics through training and professional development. To reach this point, however, future research must first investigate the effects of non-cognitive characteristics on a more representative sample of teachers. Therefore, I would like to further develop already ongoing researcher-practitioner partnerships with the University of North

Carolina General Administration (UNCGA) and the North Carolina New Teacher Support Program (NTSP).

In partnership with the teacher preparation programs at UNC system institutions, I (along with colleagues at each of these institutions) would gather measures of non-cognitive skills and traits for those individuals in their final year of teacher preparation (undergraduate or Masters of Arts in Teaching). As with previous research projects with the UNCGA, this initiative could start small, with the four or five institutions that have the greatest capacity for such work, and over time, spread system-wide. To consider a sample of teachers with greater diversity in preparation experiences (particularly alternative entry teachers), I would also collect non-cognitive measures for participants in the North Carolina NTSP, a comprehensive novice teacher induction program currently administered in the state's lowest-performing (Race to the Top) schools. The main research questions from this work would be whether non-cognitive characteristics predict teacher effectiveness, evaluation ratings, and/or retention—both overall and in teachers' original placement school. Further work on a sub-sample of these teachers could use rubric-based classroom observations, such as CLASS or the Framework for Teaching, to determine whether non-cognitive attributes predict specific teacher behaviors. Based on findings from my dissertation and comparable research initiatives (Barrick & Mount, 1991; Dobbie, 2011; Duckworth, Quinn, & Seligmann, 2009; Robertson-Kraft & Duckworth, 2013; Rockoff, Jacob, Kane, & Staiger, 2011), I believe the most promising non-cognitive attributes to initially investigate include (1) grittiness/perseverance (2) conscientiousness/organizational ability and (3) academic expectations/teacher self-

efficacy. Measures already exist for these characteristics, meaning that with the support of research partners, work could get underway quickly.

Concerning my second dissertation chapter, it is clear that, on average, teachers traditionally prepared out-of-state are less effective than in-state traditionally prepared and alternative entry teachers in elementary grades mathematics and reading. While such findings are important, especially since previous studies typically grouped all traditionally prepared teachers together, the significant advance of this work was testing three-research based hypotheses and better understanding why out-of-state prepared teachers struggle. Moving forward, the potential exists for further work on out-of-state prepared teachers,[1] however, I believe the necessary extension of this work is to question why other teacher preparation groups are less (more) effective in North Carolina public schools. Here, the natural group to consider is alternative entry teachers (excluding TFA corps members) in high schools. Findings show that alternative entry teachers (1) are significantly less effective than traditionally prepared teachers in high school science, mathematics, and social studies and (2) comprise one-half, one-third, and one-fourth of the early-career teachers in these subject areas, respectively (Henry, Bastian, Fortner, Kershaw, Purtell, Thompson, & Zulli, 2013).[2] Therefore, given the statistical and practical significance of these results, I would like to test the following explanations for alternative entry underperformance: (1) in response to lowered barriers of entry into

[1] Examples of such follow-up out-of-state studies include: (1) determining whether participation in a comprehensive induction program eases the transition for out-of-state prepared teachers into the state and benefits performance or (2) investigating whether the cognitive ability of out-of-state prepared teachers working in North Carolina differs from that of out-of-state teachers who secured employment in their preparation state.

[2] A recent report on teacher turnover issued by the North Carolina Department of Public Instruction found that school districts reported the most difficulty in filling secondary grades science and mathematics teaching positions (NCDPI, 2012).

teaching, alternative entry instructors have less human capital than traditionally prepared teachers; (2) due to lower costs of exit from the profession, alternative entry teachers attrite at significantly higher rates and withdraw job-related effort upon exiting; and (3) alternative entry teachers are less effective in their initial years in the profession—when they are concurrently taking courses and completing licensure requirements—but make steeper gains in effectiveness over time. Administrative data is available to begin testing these hypotheses; findings can shape state teacher preparation and licensure policies.

Finally, results from my last dissertation chapter indicate that characteristics of early-career principals' previous work environments, particularly the value-added of the assistant principal school (apprentice experience), are significantly associated with student achievement gains. While these results are not causal, the findings suggest that (1) early-career principals may be learning key aspects of effective school leadership during their assistant principal experiences and (2) school districts may increase principal performance through assignment patterns that maximize on-the-job learning during the apprenticeship. To advance this research and better isolate the apprenticeship learning environment effect, I would first like to conduct a non-experimental analysis with a more rigorous methodological approach—differences-in-differences with school fixed effects. Additional aspects of this research would (1) examine the process by which individuals are assigned to their assistant principal and principal schools, to determine whether there is evidence of selection or compensatory assignment and (2) test additional measures of the school environment, such as teacher responses to a working conditions survey, to verify that school value-added captures a valid construct of a high-quality learning environment. If the results from this work corroborate those in my dissertation chapter, I

believe that it would then be important to isolate specific assistant principal experiences that contribute to the development of effective school leadership practices—for assistant principals, which mastery experiences (e.g. observing teachers, planning professional development) or vicarious experiences (e.g. observing an effective principal conduct teacher evaluations) shape early-career principal performance. To complete this work—which would include primary data collection efforts (surveys and interviews)—I would partner with several school districts across the state. These partnerships may also facilitate randomized studies—randomly assigning assistant principals and principals to schools—that determine the causal nature of the apprenticeship learning environment.

Overall, completion of this dissertation has been an educative experience, further developing my (1) knowledge of the teacher and principal quality research literature (2) skill in managing administrative datasets (3) competency in considering validity threats and employing appropriate quantitative methods and (4) capacity to analyze findings and draw forth research and policy implications. The next step is building a research agenda from this foundation—with some plans outlined here—that significantly contributes to the betterment of teacher and principal quality and student academic outcomes.

APPENDIX A

ADDITIONAL RESULTS FROM CHAPTER ONE

Appendix Table A1: Teacher Evaluation Scores from the 2010-11 School Year

McREL Evaluation Ratings: All Other Teachers					
	Standard 1 Leadership	**Standard 2 Respectful Environment**	**Standard 3 Content Knowledge**	**Standard 4 Facilitate Learning**	**Standard 5 Reflect on Practice**
Distinguished	4,368	2,734	3,767	4,659	4,409
Accomplished	17,082	19,300	17,569	21,172	18,545
Proficient	22,955	22,213	23,175	18,903	21,660
Developing	1,444	1,610	1,371	1,137	1,267
Not Demonstrated	61	53	28	39	29
Observations	45,910	45,910	45,910	45,910	45,910
McREL Evaluation Ratings: TFA Corps Member Sample					
Distinguished	3	1	2	3	6
Accomplished	69	55	47	64	56
Proficient	170	183	195	175	175
Developing	9	12	7	9	14
Not Demonstrated	0	0	0	0	0
Cases	251	251	251	251	251

Note: This table displays counts for unique evaluation ratings given to teachers.

Appendix Table A2: Do Non-Cognitive Skills and Traits Predict Teacher Effectiveness in Elementary Grades

	Elementary Grades Mathematics			Elementary Grades Reading		
Factor	**1st Year Teachers**	**2nd Year Teachers**	**Combined**	**1st Year Teachers**	**2nd Year Teachers**	**Combined**
Factor 1: Achieve and Critical Think	-0.018 (0.023)	-0.014 (0.043)	-0.007 (0.022)	0.013 (0.017)	-0.051 (0.044)	-0.003 (0.016)
Factor 2: Leadership and perseverance	-0.006 (0.029)	-0.007 (0.043)	-0.009 (0.025)	-0.025 (0.019)	-0.017 (0.035)	-0.008 (0.017)
Factor 3: Organ. and Motiv	**0.049**$^{+}$ (0.025)	**0.085*** (0.037)	**0.054*** (0.022)	**0.053*** (0.021)	**0.080**$^{+}$ (0.044)	**0.047**** (0.016)
Factor 4: Respect and Fit	0.014 (0.020)	-0.008 (0.036)	0.008 (0.019)	0.008 (0.020)	0.025 (0.034)	0.010 (0.017)
Cases	3,072	1,599	4,671	4,024	2,301	6,325

Note: + Indicates significance at the p<0.10 level; * indicates significance at the p<0.05 level; ** indicates significance at the p<0.01 level.

Appendix Table A3: Do Non-Cognitive Skills and Traits Predict Teacher Effectiveness in High School?

	All High School End of Course Exams			STEM Subjects			Non-STEM Subjects		
Factor	**1st Year Teachers**	**2nd Year Teachers**	**Combined**	**1st Year Teachers**	**2nd Year Teachers**	**Combined**	**1st Year Teachers**	**2nd Year Teachers**	**Combined**
Factor 1: Ach and Critical	-0.033 (0.022)	0.014 (0.025)	-0.017 (0.021)	-0.033 (0.025)	0.042 (0.029)	-0.004 (0.025)	-0.035 (0.026)	-0.021 (0.035)	**-0.062*** (0.024)
Factor 2: Lead and Pers	-0.009 (0.023)	**-0.074*** (0.029)	-0.023 (0.023)	-0.005 (0.033)	**-0.079**** (0.027)	-0.026 (0.028)	-0.013 (0.021)	-0.092 (0.057)	-0.027 (0.022)
Factor 3: Organ and Motiv	0.016 (0.019)	**0.049*** (0.022)	0.023 (0.016)	0.017 (0.024)	**0.047**+ (0.028)	0.024 (0.021)	-0.034 (0.023)	0.025 (0.039)	-0.015 (0.021)
Factor 4: Respect and Fit	**0.044**+ (0.026)	0.014 (0.022)	0.032 (0.021)	0.026 (0.029)	0.008 (0.030)	0.021 (0.025)	**0.058**+ (0.031)	**0.087*** (0.036)	**0.073**** (0.024)
Cases	10,409	7,145	17,554	7,287	5,029	12,316	3,122	2,116	5,238

Note: + Indicates significance at the p<0.10 level; * indicates significance at the p<0.05 level; ** indicates significance at the p<0.01 level.

Appendix Table A4: Do Non-Cognitive Skills & Traits Predict Teacher Evaluation Ratings?

Factor	Standard 1: Leadership	Standard 2: Respectful Environment	Standard 3: Content Knowledge	Standard 4: Facilitate Learning	Standard 5: Reflect on Practice
Factor 1: Ach and Critical	1.138 (0.66)	**1.668*** (2.31)	1.233 (1.04)	**1.727**** (2.65)	1.294 (1.16)
Factor 2: Lead and Pers	**1.444*** (2.08)	**1.381**$^{+}$ (1.95)	1.046 (0.22)	1.256 (1.11)	0.963 (-0.19)
Factor 3: Organ. and Motiv	1.176 (0.95)	**1.520*** (2.19)	**1.500**$^{+}$ (1.90)	1.241 (1.20)	1.240 (1.13)
Factor 4: Respect and Fit	1.011 (0.06)	0.836 (-1.01)	0.795 (-1.27)	**0.714**$^{+}$ (-1.95)	0.923 (-0.42)
Cases	249	249	249	249	249

Note: Models include teacher experience controls and a rich set of school characteristics. The sample includes all corps members evaluated by their principal in 2010-11. Cells report odds ratios for being rated *above proficient* and z-scores. + indicates significance at the p<0.10 level; * indicates significance at the p<0.05 level; ** indicates significance at the p<0.01 level.

APPENDIX B

ESTIMATING THE APPRENTICESHIP LEARNING ENVIRONMENT

For this analysis we quantify an assistant principal's learning environment quality using a measure of school value-added. Due to censoring concerns—some individuals are assistant principals for longer periods than we can examine the value-added learning environment—and the hypothesis that the most recent assistant principal experiences exert the strongest effect on early-career principal effectiveness, we focus on the apprenticeship learning environment quality from the year immediately prior to entrance into the principalship. To estimate this value any specification needed to address the following two objectives: (1) it must measure school (not principal) effectiveness, since we want to identify the overall learning environment an assistant principal experienced and (2) it must generate yearly school value-added estimates for each North Carolina public school from 2005-06 through 2008-09—the four year period in which our sample of first-time principals last worked as assistant principals.

To address the requirements above, we estimated the quality of the apprenticeship learning environment using a two-level random effects model with a rich set of student and school covariates.[1] Here, we identified school-year value-added with the school-level random effect, which represents the unexplained variation in achievement between schools. Given the rich set of covariates employed in this measurement model (the same student and school covariates as shown in Table 3.1), we argue that this school-year residual is attributable to the actions of school leadership and teachers, capturing the

[1] Currently, there is no agreement on whether to control for teacher and classroom characteristics when estimating school effectiveness. Following Grissom and colleagues we exclude these covariates from our school value-added specifications (Grissom, Kalogrides, & Loeb, 2012).

apprenticeship learning environment quality. Equations for the school-year random effect and the full measurement model are as follows:

$r_s = Y - \hat{Y}$ (1)

where r_s represents the school-year residual measuring school effectiveness;

Y is the actual test score outcome for the school;

and $\hat{Y}$ is the predicted test score outcome for the school, given the rich set of controls.

$Y_{ist} = \beta_0 + \beta_1 Y_{it-n} + \gamma_x X_{ist} + \gamma_w W_{st} + \varepsilon_i + r_s$ (2)

where Y_{ist} is the test score for student *i* in school *s* at time *t*;

Y_{it-n} represents the prior test score(s) for student *i*;

X_{ist} and W_{st} represent the set of individual student and school covariates;

γ_x and γ_w represent a sector of fixed, average effects for each student and school covariate;

and ε_i and r_s are terms representing the unexplained variation at the student and school levels, respectively.

For these models we ran separate analyses in elementary grades mathematics and reading, middle grades mathematics and reading, and a combined model for the five high school EOC exams required for graduation—algebra 1, English 1, biology, civics, and U.S. history. We included all North Carolina public schools in these specifications, rather than limit the data to the schools our sample of first-time principals apprenticed in, to take advantage of the full range of variability in the data and to create focal variables based on the full population of schools. Post-estimation, we averaged the math and reading value-added random effects for elementary and middle grades to generate a single school-year measure of the apprenticeship learning environment quality. Finally,

with this measure of the apprenticeship learning environment quality we created two focal variables for this analysis: (1) a standardized, continuous measure of the school effectiveness and (2) a dichotomous indicator for being an assistant principal in a top quintile value-added school.

APPENDIX C

ADDITIONAL RESULTS FROM CHAPTER THREE

Appendix Table C1: Individual Principal Characteristics and Student Achievement (Full Sample)

	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
Focal Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Female Principal	-0.007 (0.007)	-0.007 (0.007)	0.009 (0.005)	0.002 (0.005)	0.000 (0.007)	-0.003 (0.008)	0.006 (0.004)	0.005 (0.004)	0.008 (0.013)	-0.015 (0.015)
Cases	566,293		770,983		423,328		446,125		428,450	
Minority Principal	-0.009 (0.008)	-0.007 (0.008)	-0.001 (0.007)	0.001 (0.006)	**-0.029**** (0.009)	-0.016 (0.010)	0.002 (0.006)	-0.004 (0.006)	0.013 (0.022)	-0.008 (0.022)
Cases	564,601		768,683		423,328		446,125		426,884	
Selectivity										
Competitive	-0.003 (0.009)	-0.012 (0.008)	0.003 (0.007)	0.011 (0.007)	-0.009 (0.010)	**-0.033**** (0.012)	-0.001 (0.006)	**-0.016*** (0.007)	**-0.056**** (0.016)	-0.024 (0.020)
Highly Comp	0.001 (0.010)	-0.011 (0.008)	-0.001 (0.007)	0.014 (0.007)	-0.021 (0.012)	-0.024 (0.014)	-0.009 (0.006)	-0.011 (0.008)	**-0.070**** (0.019)	-0.026 (0.020)
Cases	566,293		770,983		423,328		446,125		428,450	
NBC	**0.021*** (0.010)	0.007 (0.010)	**0.017*** (0.008)	0.006 (0.008)	-0.022 (0.015)	-0.013 (0.016)	**-0.013*** (0.006)	-0.012 (0.009)	0.002 (0.026)	-0.015 (0.031)
Cases	566,293		770,983		423,328		446,125		428,450	
Std. Principal Exam Scores	**0.012**** (0.004)	0.006 (0.004)	0.003 (0.003)	0.000 (0.003)	**0.009*** (0.004)	-0.000 (0.005)	0.002 (0.003)	0.004 (0.003)	0.002 (0.006)	-0.017 (0.009)
Cases	538,402		731,710		393,767		414,060		408,354	
Top Quintile Exam Scores	**0.018*** (0.008)	0.010 (0.008)	**0.015*** (0.006)	0.003 (0.006)	**0.019*** (0.009)	0.011 (0.010)	0.002 (0.005)	0.007 (0.006)	0.001 (0.013)	**-0.052**** (0.019)
Cases	538,402		731,710		393,767		414,060		408,354	

Appendix Table C1 (Continued): Individual Principal Characteristics and Student Achievement (Full Sample)

	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
Focal Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
NC Principal Fellow	**-0.022*** (0.009)	**-0.023*** (0.009)	**-0.022**** (0.006)	-0.009 (0.007)	0.001 (0.016)	-0.029 (0.017)	-0.000 (0.007)	-0.009 (0.009)	-0.002 (0.021)	0.021 (0.037)
Cases	566,293		770,983		423,328		446,125		428,450	
Principal Prep										
Private Masters	0.003 (0.012)	0.007 (0.013)	-0.004 (0.008)	-0.006 (0.009)	-0.014 (0.012)	0.005 (0.015)	-0.003 (0.006)	0.011 (0.008)	-0.005 (0.017)	-0.032 (0.020)
Out-of-state Masters	0.001 (0.011)	-0.006 (0.011)	0.007 (0.009)	0.004 (0.009)	0.017 (0.013)	**0.036**** (0.013)	0.013 (0.007)	**0.018*** (0.007)	-0.010 (0.024)	0.005 (0.021)
Public Doctorate	0.006 (0.013)	0.003 (0.013)	**-0.033**** (0.009)	-0.008 (0.009)	**-0.025*** (0.012)	-0.016 (0.014)	-0.012 (0.008)	0.009 (0.009)	0.041 (0.028)	-0.023 (0.029)
Private Doctorate	-0.033 (0.071)	0.003 (0.038)	-0.072 (0.048)	**-0.114**** (0.034)	0.044 (0.039)	---	**-0.072**** (0.018)	---	0.014 (0.029)	-0.051 (0.037)
Out-of-State Doctorate	**-0.033*** (0.016)	-0.013 (0.019)	-0.001 (0.014)	-0.007 (0.014)	**-0.036*** (0.017)	-0.010 (0.027)	0.013 (0.008)	-0.004 (0.008)	0.012 (0.026)	-0.049 (0.026)
Cases	565,246		769,921		423,328		446,125		428,450	

Note: Odd numbered columns display results from rich covariate adjustment models (OLS) with teacher/classroom covariates. Even numbered columns display results from models with school district fixed effects and teacher/classroom covariates. * Indicates statistical significance at the p<0.05 level; ** indicates statistical significance at the p<0.01 level.

Appendix Table C2: Previous Work Environment Characteristics and Student Achievement (Full Sample)

	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
Focal Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Teacher-School Match	**-0.032***	**-0.048****	0.005	0.003	-0.004	0.000	-0.004	-0.013	0.004	-0.006
	(0.012)	(0.012)	(0.010)	(0.009)	(0.011)	(0.011)	(0.006)	(0.007)	(0.017)	(0.021)
Cases	566,293		770,983		423,328		446,125		428,450	
AP-School Match	0.004	0.008	0.006	0.001	0.013	**0.020***	**0.013****	0.008	0.009	0.019
	(0.008)	(0.008)	(0.006)	(0.006)	(0.008)	(0.010)	(0.004)	(0.005)	(0.014)	(0.015)
Cases	566,293		770,983		423,328		446,125		428,450	
AP-School Type Match	-0.009	-0.014	-0.009	-0.003	**0.018***	**0.019***	**0.008***	**0.012***	**0.068****	**0.042***
	(0.007)	(0.008)	(0.006)	(0.006)	(0.008)	(0.008)	(0.004)	(0.005)	(0.019)	(0.018)
Cases	566,293		770,983		423,328		446,125		428,450	
Std. Apprenticeship Learning Environment	**0.017****	**0.011***	**0.011****	0.004	**0.023****	**0.015****	**0.009****	0.004	**0.032****	**0.023****
	(0.004)	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)	(0.002)	(0.003)	(0.007)	(0.007)
Cases	518,843		704,197		403,579		425,511		401,981	
Top Quintile Apprenticeship Learning Environment	**0.023****	0.008	**0.019****	**0.013***	**0.040****	**0.036****	**0.018****	**0.020****	**0.062****	**0.060****
	(0.009)	(0.009)	(0.006)	(0.006)	(0.009)	(0.010)	(0.005)	(0.005)	(0.016)	(0.015)
Cases	518,843		704,197		403,579		425,511		401,981	

Note: Odd numbered columns display results from rich covariate adjustment models (OLS) with teacher/classroom covariates. Even numbered columns display results from models with school district fixed effects and teacher/classroom covariates. * Indicates statistical significance at the p<0.05 level; ** indicates statistical significance at the p<0.01 level.

Appendix Table C3: Individual Principal Characteristics and Student Achievement (1st & 2nd Year)

	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
Focal Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Female Principal	-0.001 (0.008)	0.002 (0.008)	**0.016*** (0.006)	0.012 (0.006)	-0.007 (0.010)	-0.004 (0.010)	0.001 (0.006)	-0.003 (0.006)	0.025 (0.017)	-0.018 (0.018)
Cases	374,403		525,763		239,983		252,061		239,123	
Minority Principal	-0.007 (0.010)	-0.008 (0.010)	0.002 (0.008)	0.001 (0.008)	**-0.034**** (0.012)	-0.010 (0.011)	-0.003 (0.008)	-0.006 (0.007)	0.047 (0.030)	0.029 (0.024)
Cases	373,050		523,802		239,983		252,061		238,120	
Selectivity										
Competitive	-0.000 (0.011)	-0.015 (0.010)	-0.006 (0.008)	0.005 (0.008)	0.013 (0.013)	-0.003 (0.016)	0.002 (0.008)	-0.012 (0.010)	**-0.076**** (0.024)	-0.024 (0.026)
Highly Comp	0.009 (0.013)	-0.010 (0.012)	-0.012 (0.009)	0.008 (0.009)	-0.003 (0.016)	0.002 (0.020)	-0.007 (0.009)	0.000 (0.011)	**-0.090**** (0.025)	-0.015 (0.024)
Cases	374,403		525,763		239,983		252,061		239,123	
NBC	**0.029*** (0.013)	0.015 (0.012)	0.016 (0.009)	0.009 (0.009)	**-0.030*** (0.014)	**-0.033*** (0.015)	-0.016 (0.009)	**-0.024*** (0.012)	0.007 (0.033)	0.040 (0.043)
Cases	374,403		525,763		239,983		252,061		239,123	
Std. Principal Exam Scores	**0.016**** (0.005)	**0.012*** (0.005)	0.005 (0.004)	0.004 (0.003)	0.008 (0.006)	-0.003 (0.006)	-0.000 (0.003)	0.000 (0.004)	-0.006 (0.009)	-0.016 (0.011)
Cases	355,740		499,117		222,724		233,141		224,315	
Top Quintile Exam Scores	**0.022*** (0.010)	0.016 (0.010)	**0.022**** (0.008)	0.011 (0.007)	0.008 (0.010)	-0.003 (0.012)	-0.006 (0.007)	-0.000 (0.008)	-0.001 (0.019)	**-0.061*** (0.026)
Cases	355,740		499,117		222,724		233,141		224,315	

Appendix Table C3 (Continued): Individual Principal Characteristics and Student Achievement (1st & 2nd Year)

	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
Focal Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
NC Principal Fellow	-0.015	-0.014	**-0.026****	-0.011	-0.003	-0.037	-0.005	**-0.026***	0.003	-0.028
	(0.012)	(0.012)	(0.008)	(0.008)	(0.022)	(0.022)	(0.009)	(0.011)	(0.034)	(0.044)
Cases	374,403		525,763		239,983		252,061		239,123	
Principal Prep										
Private Masters	0.023	0.026	0.005	0.008	0.005	0.029	0.002	**0.025***	-0.007	-0.022
	(0.014)	(0.015)	(0.010)	(0.011)	(0.014)	(0.018)	(0.009)	(0.012)	(0.020)	(0.026)
Out-of-state Masters	0.002	0.001	0.006	0.005	0.016	**0.037***	0.014	0.010	-0.007	0.044
	(0.013)	(0.014)	(0.010)	(0.011)	(0.017)	(0.017)	(0.008)	(0.009)	(0.028)	(0.030)
Public Doctorate	0.011	0.011	**-0.033****	-0.003	-0.021	-0.001	-0.004	0.021	0.067	0.008
	(0.017)	(0.016)	(0.012)	(0.011)	(0.016)	(0.017)	(0.012)	(0.012)	(0.035)	(0.041)
Private Doctorate	-0.068	-0.022	**-0.114****	**-0.135****	0.095	---	**-0.052***	---	0.019	0.027
	(0.086)	(0.035)	(0.039)	(0.031)	(0.054)		(0.020)		(0.043)	(0.049)
Out-of-State Doctorate	-0.041	-0.016	-0.011	-0.008	-0.046	0.012	-0.009	-0.008	0.018	0.010
	(0.021)	(0.022)	(0.016)	(0.015)	(0.028)	(0.030)	(0.011)	(0.013)	(0.035)	(0.030)
Cases	373,914		525,277		239,983		252,061		239,123	

Note: Odd numbered columns display results from rich covariate adjustment models (OLS) without teacher/classroom covariates. Even numbered columns display results from models with school district fixed effects and no teacher/classroom covariates. * Indicates statistical significance at the p<0.05 level; ** indicates statistical significance at the p<0.01 level.

Appendix Table C4: Individual Principal Characteristics and Student Achievement (3rd-5th Year)

	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
Focal Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Female Principal	-0.021	**-0.026***	-0.005	**-0.018***	0.010	-0.005	0.011	0.008	0.008	-0.021
	(0.012)	(0.012)	(0.009)	(0.008)	(0.011)	(0.014)	(0.006)	(0.006)	(0.018)	(0.019)
Cases	191,890		245,220		183,345		194,064		189,327	
Minority Principal	-0.018	-0.009	-0.002	0.002	-0.026	-0.020	0.002	-0.012	-0.033	**-0.061***
	(0.014)	(0.013)	(0.011)	(0.011)	(0.014)	(0.018)	(0.008)	(0.009)	(0.027)	(0.027)
Cases	191,551		244,881		183,345		194,064		188,764	
Selectivity										
Competitive	-0.009	-0.008	0.013	**0.024***	-0.030	**-0.068****	-0.008	-0.018	**-0.053***	-0.054
	(0.013)	(0.014)	(0.011)	(0.012)	(0.016)	(0.018)	(0.008)	(0.010)	(0.023)	(0.030)
Highly Comp	-0.009	-0.013	0.016	0.024	-0.030	**-0.056****	-0.011	-0.010	**-0.072****	-0.065
	(0.015)	(0.015)	(0.012)	(0.013)	(0.017)	(0.019)	(0.009)	(0.011)	(0.029)	(0.035)
Cases	191,890		245,220		183,345		194,064		189,327	
NBC	0.003	-0.009	0.018	0.003	-0.011	-0.005	-0.004	-0.004	-0.012	-0.052
	(0.018)	(0.017)	(0.016)	(0.015)	(0.023)	(0.025)	(0.009)	(0.011)	(0.037)	(0.041)
Cases	191,890		245,220		183,345		194,064		189,327	
Std. Principal Exam Scores	0.005	-0.004	-0.001	-0.003	0.012	0.004	0.006	**0.008***	0.010	0.001
	(0.006)	(0.007)	(0.005)	(0.005)	(0.007)	(0.009)	(0.004)	(0.004)	(0.009)	(0.011)
Cases	182,662		232,593		171,043		180,919		184,039	
Top Quintile Exam Scores	0.006	-0.006	-0.005	-0.013	**0.037****	**0.036***	0.010	**0.015***	0.005	-0.020
	(0.013)	(0.014)	(0.009)	(0.009)	(0.013)	(0.017)	(0.007)	(0.007)	(0.019)	(0.023)
Cases	182,662		232,593		171,043		180,919		184,039	

Appendix Table C4 (Continued): Individual Principal Characteristics and Student Achievement (3rd-5th Year)

	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
Focal Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
NC Principal Fellow	**-0.038***	**-0.033***	-0.014	-0.001	0.013	-0.021	0.007	0.005	0.007	0.011
	(0.015)	(0.014)	(0.011)	(0.012)	(0.021)	(0.023)	(0.011)	(0.013)	(0.031)	(0.034)
Cases	191,890		245,220		183,345		194,064		189,327	
Principal Prep										
Private Masters	-0.033	-0.040	-0.021	**-0.036***	-0.026	-0.015	-0.010	-0.002	0.005	-0.050
	(0.018)	(0.023)	(0.012)	(0.016)	(0.019)	(0.025)	(0.008)	(0.010)	(0.031)	(0.030)
Out-of-state Masters	0.000	-0.022	0.011	-0.003	0.009	0.030	0.011	**0.024***	-0.014	-0.014
	(0.018)	(0.019)	(0.016)	(0.014)	(0.017)	(0.021)	(0.010)	(0.011)	(0.035)	(0.031)
Public Doctorate	0.008	-0.012	-0.029	-0.012	-0.026	-0.035	-0.018	0.004	0.046	**-0.115****
	(0.020)	(0.021)	(0.015)	(0.014)	(0.018)	(0.020)	(0.010)	(0.010)	(0.041)	(0.031)
Private Doctorate	0.042	---	0.033	---	-0.034	---	**-0.051****	---	-0.009	**-0.210****
	(0.070)		(0.078)		(0.050)		(0.010)		(0.036)	(0.033)
Out-of-State Doctorate	-0.019	0.006	0.013	0.007	**-0.040***	-0.006	**0.021***	-0.000	0.014	**-0.150****
	(0.022)	(0.026)	(0.023)	(0.029)	(0.020)	(0.035)	(0.009)	(0.012)	(0.039)	(0.026)
Cases	191,332		244,644		183,345		194,064		189,327	

Note: Odd numbered columns display results from rich covariate adjustment models (OLS) without teacher/classroom covariates. Even numbered columns display results from models with school district fixed effects and no teacher/classroom covariates. * Indicates statistical significance at the p<0.05 level; ** indicates statistical significance at the p<0.01 level.

Appendix Table C5: Previous Work Environment Characteristics and Student Achievement (1st and 2nd Year)

	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
Focal Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Teacher-School Match	**-0.045***	**-0.052****	-0.008	-0.013	0.015	0.010	0.006	-0.011	0.013	0.027
	(0.017)	(0.015)	(0.013)	(0.011)	(0.015)	(0.013)	(0.009)	(0.010)	(0.027)	(0.028)
Cases	374,403		525,763		239,983		252,061		239,123	
AP-School Match	0.012	0.018	0.013	0.005	0.007	0.020	0.011	0.008	0.016	0.029
	(0.010)	(0.010)	(0.008)	(0.007)	(0.011)	(0.012)	(0.006)	(0.007)	(0.019)	(0.022)
Cases	374,403		525,763		239,983		252,061		239,123	
AP-School Type Match	-0.011	-0.015	-0.006	0.000	0.012	0.011	0.003	0.004	**0.119****	**0.072****
	(0.009)	(0.009)	(0.007)	(0.007)	(0.011)	(0.011)	(0.006)	(0.006)	(0.026)	(0.024)
Cases	374,403		525,763		239,983		252,061		239,123	
Std. Apprenticeship Learning Environment	**0.018****	**0.011***	**0.008***	0.003	**0.024****	**0.019****	**0.015****	**0.010****	**0.041****	**0.036****
	(0.005)	(0.005)	(0.004)	(0.003)	(0.006)	(0.006)	(0.003)	(0.003)	(0.011)	(0.008)
Cases	338,451		475,010		227,679		239,087		222,609	
Top Quintile Apprenticeship Learning Environment	**0.026***	0.012	**0.018***	**0.017***	**0.055****	**0.054****	**0.033****	**0.032****	**0.084****	**0.080****
	(0.010)	(0.010)	(0.008)	(0.008)	(0.010)	(0.012)	(0.007)	(0.007)	(0.023)	(0.019)
Cases	338,451		475,010		227,679		239,087		222,609	

Note: Odd numbered columns display results from rich covariate adjustment models (OLS) without teacher/classroom covariates. Even numbered columns display results from models with school district fixed effects and no teacher/classroom covariates. * Indicates statistical significance at the p<0.05 level; ** indicates statistical significance at the p<0.01 level.

Appendix Table C6: Previous Work Environment Characteristics and Student Achievement (3rd-5th Year)

	Elementary Math		Elementary Reading		Middle Math		Middle Reading		High School	
Focal Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Teacher-School Match	-0.010	**-0.054****	0.022	**0.026***	-0.016	0.000	-0.014	**-0.023***	-0.021	**-0.080****
	(0.017)	(0.017)	(0.014)	(0.013)	(0.017)	(0.022)	(0.009)	(0.009)	(0.022)	(0.030)
Cases	191,890		245,220		183,345		194,064		189,327	
AP-School Match	-0.005	-0.003	-0.009	-0.002	**0.030***	**0.045****	**0.015***	0.003	-0.001	0.008
	(0.013)	(0.013)	(0.010)	(0.009)	(0.012)	(0.016)	(0.006)	(0.008)	(0.020)	(0.025)
Cases	191,890		245,220		183,345		194,064		189,327	
AP-School Type Match	-0.008	-0.006	-0.019	-0.015	**0.031****	**0.041****	0.011	**0.016***	0.016	**0.059***
	(0.013)	(0.015)	(0.010)	(0.010)	(0.011)	(0.011)	(0.006)	(0.007)	(0.025)	(0.026)
Cases	191,890		245,220		183,345		194,064		189,327	
Std. Apprenticeship Learning Environment	**0.017***	**0.014***	**0.018****	0.009	**0.019****	0.009	0.001	-0.004	0.017	0.006
	(0.007)	(0.007)	(0.005)	(0.005)	(0.006)	(0.008)	(0.003)	(0.004)	(0.009)	(0.009)
Cases	180,392		229,187		175,900		186,424		179,372	
Top Quintile Apprenticeship Learning Environment	0.017	-0.002	**0.028****	0.010	0.016	0.012	-0.002	0.008	**0.048***	**0.046***
	(0.015)	(0.015)	(0.011)	(0.010)	(0.014)	(0.018)	(0.008)	(0.008)	(0.021)	(0.021)
Cases	180,392		229,187		175,900		186,424		179,372	

Note: Odd numbered columns display results from rich covariate adjustment models (OLS) without teacher/classroom covariates. Even numbered columns display results from models with school district fixed effects and no teacher/classroom covariates. * Indicates statistical significance at the p<0.05 level; ** indicates statistical significance at the p<0.01 level.

REFERENCES

Barrick, M., & Mount, M. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*(1), 1-26.

Dobbie, W. (2011). Teacher characteristics and student achievement: Evidence from Teach For America. Available from: http://www.people.fas.harvard.edu.

Duckworth, A., Quinn, P., & Seligman, M. (2009). Positive predictors of teacher effectiveness. *The Journal of Positive Psychology, 4*(6), 540-547.

Henry, G.T., Bastian, K.C., Fortner, C.K., Kershaw, D.C., Purtell, K.M., Thompson, C.L., & Zulli, R.A. (2013). Teacher preparation policies and their effects on student achievement. In press, *Education Finance and Policy*.

North Carolina Department of Public Instruction. (2012). Teacher turnover report: Annual report on the reasons teachers leave 2011-2012. Available from: http://www.ncpublicschools.org/docs/recruitment/surveys/turnover/2011-12turnoverreport.pdf.

Robertson, Kraft, C., & Duckworth, A. (2013). True grit: Trait-level perseverance and passion for long-term goals predicts effectiveness and retention among novice teachers. Available from: http://www.sas.upenn.edu/~duckwort/images/publications/truegrit.pdf.

Rockoff, J., Jacob, B., Kane, T., & Staiger, D. (2011). Can you recognize an effective teacher when your recruit one? *Education Finance and Policy, 6*(1), 43-74.