Assistant Principal or Assistant to the Principal?
Exploring the Contributions, Duties, and Preparation of Assistant Principals

By
David S. Woo

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
in
Leadership and Policy Studies
June 30, 2020
Nashville, Tennessee

Approved:
Jason A. Grissom, Ph.D.
Ellen B. Goldrin, Ph.D.
Marisa A. Cannata, Ph.D.
Andrea K. Rorrer, Ph.D.

To Grace, whose labor produced this work as much as mine,
and
Joanna, who inspires me to work towards a better future

ACKNOWLEDGMENTS

This dissertation was a product of love and sacrifice on the part of many. My time as a doctoral student has been challenging, and I would not have been able to complete this project if it were not for my family, friends, and mentors pouring into me. I am so grateful for the many that sacrificed of their time and energy to help me reach this point. In this space, I would like to name a few people who have been particularly instrumental in completing this project.

First, I want to thank the Tennessee Education Research Alliance and Tennessee Department of Education who provided the data for this project. This project would not have been possible without their support.

The members of my dissertation committee who provided thoughtful critiques and constant support in a strange academic year. Ellen Goldring, Andrea Rorer, and Marisa Cannata are inspiring, brilliant women, and I am a better scholar because I have known them.

I am thankful for my dissertation chair and advisor, Jason Grissom, who has been tremendously patient and supportive throughout my doctoral education but especially in this last year. He has opened countless doors to opportunities and pushed me to always produce my best work. I have learned so much from watching and working with him these past five years, and I cannot begin to put into words what his mentorship has meant to me.

My friends and colleagues in LPO who I have commiserated and celebrated with along this journey. They are all amazing people, but I am thankful for my cohort most of all. Its members/My classmates inspire me to work harder and think deeper.

The teachers I worked with in Chicago taught me so much about the power of education. The care and diligence with which they approached their work demonstrated to me that a great

educator can have a lasting impact on students' lives. I am also grateful for the students who were in my classrooms. They continue to remind me of the stakes of education research and policy, and their resilience in the face of systems that oppress them reminds me I can still do more.

My parents instilled in me the value of education and they encouraged me to make a difference in whatever I did. Their support and sacrifice as immigrants have made my academic journey possible. My daughter, Joanna, who was born in the midst of my dissertating, reminded me what is most important in life.

The person I want to thank most is my partner, Grace. She has had to endure the worst versions of my self, yet was the first to celebrate my accomplishments. She went to incredible lengths to support me, all while working a full-time job, caring for me while I recovered from a motorcycle accident, giving birth to our daughter, and obtaining her own graduate degree. This work is unequivocally hers as much as it is mine.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS ..... **iv**

LIST OF FIGURES ..... **viii**

LIST OF TABLES ..... **ix**

INTRODUCTION **1**

Chapter

I. Addition by Division or Quality over Quantity: Estimating the contribution of assistant principals to school climate **4**

I.1 Literature Review 8
I.1.1 School climate 8
I.1.2 Distributed leadership 10
I.1.3 Research on APs 12
I.1.4 School climate, distributed leadership, and the role of the AP 16
I.2 APs and the Tennessee context 18
I.3 Data and Measures 22
I.2.1 Measures of exposure to APs 24
I.2.2 Outcome Measures 26
I.4 Methods 29
I.5 Findings 32
I.6 Sensitivity Analyses 36
I.7 Discussion 41
I.8 Limitations 42
I.9 Conclusion 44
Chapter I Figures 46
Chapter I Tables 51
Chapter I Appendix Figures 68

II. Determinants of Assistant Principals' Work: School organizational factors and their relationship to assistant principals' time use **79**

II.1 Literature Review 82
II.1.1 How school personnel allocate their time 82
II.1.2 What factors are related to how school personnel allocate their time 84
II.1.3 How APs allocate their time 87
II.1.4 What factors are related to how APs allocate their time 88
II.1.5 Leadership substitutes and APs' time use 89
II.1.6 Contributions 91
II.2 The Tennessee Contexts and APs' Time Allocation 91

II.3 Data and Measures 93
II.3.1 Measures capturing the time allocation of APs 96
II.3.2 Measures of support staff in schools 99
II.4 Methods 100
II.4.1 Descriptive Analyses 100
II.4.2 Regression Analyses 100
II.5 Findings 101
II.5.1 Descriptive analyses 101
II.5.2 Regression Analyses 104
II.6 Discussion 117
II.7 Limitations 119
II.7.1 Threats to internal validity 119
II.7.2 Generalizability 121
II.8 Conclusion 121
Chapter II Figures 124
Chapter II Tables 128
Chapter II Appendix Tables 141

III. Preparing the Assistant Principal: Preparation program features, perceptions, and effectiveness on the job 149

III.1 Literature Review 152
III.1.1 Effective preparation of aspiring teachers 154
III.1.2 Overview of issues and approaches to identifying effective SLPPs 156
III.1.3 Case Studies of Exemplary SLPPs 157
III.1.4 Quality features of SLPPs as identified by graduates' outcomes 159
III.1.5 Effective SLPPs as identified by graduates' outcomes 161
III.1.6 Preparing APs 164
III.1.7 Gaps in the Literature and Contributions 165
III.2 Tennessee Policy Context 166
III.3 Data and Measures 169
III.4 Methods 174
III.4.1 Coding of program features 174
III.4.2 Descriptive Analyses 177
III.4.3 Regression Analyses 177
III.5 Findings 178
III.5.1 Descriptive Analyses 178
III.5.2 Regression Analyses 181
III.6 Discussion 188
III.7 Limitations 190
III.8 Conclusion 192
Chapter III Figures 194
Chapter III Tables 197
Chapter III Appendix Figures 209
Chapter III Appendix Tables 210

REFERENCES 217

LIST OF FIGURES

Figure Page

I.1: Conceptual framework 46

I.2: Average teacher outcomes before and after a change in having an AP 47

I.3: Average student outcomes before and after a change in having an AP 48

I.4: Coefficients for teacher outcomes regressed on relative time 49

I.5: Coefficients for student outcomes regressed on relative time 50

I.A1: BEP funding formula for APs in Tennessee 68

I.A2: Knoxville County Schools' adjustments to state allocation formula 69

I.A3: Metro Nashville Public Schools' budget allocation formula 70

I.A4: Shelby County Schools' budget allocation formula 71

I.A5 Scree plot of eigenvalues for factor analysis conducted for each year 72

II.1: Percent of APs' time spent on leadership tasks over time 124

II.2: WSS, log(WSS), η 2 , and PRE for all K cluster solutions 125

II.3: Time use by position 126

II.4: Time use by AP time use profile 127

III.1: New APs' and new principals' average years since graduation 194

III.2: Proportion of SLPP graduates in education positions 195

III.3: Predicted evaluation ratings as an AP for four graduates from hypothetical SLPPs 196

III.A1 Map of SLPP main campuses in Tennessee 209

LIST OF TABLES

Table Page

I.1 Counts of schools with APs and schools that are funded for APs.....51

I.2: Schools with an AP compared to schools without an AP 52

I.3: Schools that experience a change in AP exposure as compare to schools that do not 53

I.4: Schools that gain an AP compared to schools that lose an AP 54

I.5: Schools with one AP compared to schools with multiple APs 55

I.6: Schools with low rated APs as compared to schools with high rated APs.....56

I.7: Distribution of outcomes for the three primary samples used in this study 57

I.8: Comparison of models estimating teachers' perceptions of relationships on having an AP...58

I.9: Relationship between outcomes associated with school climate and having an AP 59

I.10: Relationship between outcomes associated with school climate and the number of APs.....60

I.11: Relationship between outcomes associated with school climate and evaluation ratings of APs.....61

I.12. Sub-group analyses of student outcomes and having an AP by student race.....62

I.13. Sub-group analyses of student outcomes and number of APs by student race 63

I.14. Sub-group analyses of student outcomes and AP evaluation ratings by student race.....64

I.15: Sensitivity analyses – subgroup analyses by grade level for having an AP.....65

I.16: Sensitivity analyses – subgroup analyses by grade level for number of APs 66

I.17: Sensitivity analyses – subgroup analyses by grade level for AP evaluation rating.....67

I.A1: Survey questions and years included.....73

I.A2: Instrumental variables analysis – Instrumenting AP FTEs under BEP for number of APs .74

I.A3: Reliability measures of teachers' perceptions of relationships 75
I.A4: Sub-group analyses of student outcomes and having an AP by student race 76
I.A5: Sub-group analyses of student outcomes and AP evaluation ratings by student race 77
I.A6: Sub-group analyses of student outcomes and number of APs by student race 78
II.1: Characteristics of sample APs as compared to all other APs.....128
II.2: AP time use by clusters of time use 129
II.3: Characteristics of APs and their schools by time use cluster.....130
II.4: Time use of APs by grade level 131
II.5: Time use of APs by locale type.....132
I.6: Predicting percent of time on duties from characteristics of APs and their schools 133
II.7: Predicting AP time use profile from characteristics of APs and their schools 134
II.8: Predicting percent of time on duties from support staff in the school 135
II.9: Predicting AP time use profile from support staff in the school.....136
II.10: Predicting percent of time on duties from principal characteristics.....137
II.11: Predicting percent of time use profile from principal characteristics 138
II.12: Predicting percent of time on major duties from other leaders' percent of time on duties139
II.13: Predicting percent of time on minor duties from other leaders' percent of time on duties140
II.A1: AP counts and response rates in TES by year.....141
II.A2: Duties surveyed by year of TES.....142
II.A3: List of common position titles for support staff.....143
II.A4: Reliability coefficients for items in the scale for instructional leadership.....144
II.A5: Correlation matrix of percent of time principals allocate to duties.....145
II.A6: Correlation matrix of percent of time other APs allocate to duties.....146

II.A7: Correlation matrix of percent of time other APs and principals allocate to duties 147
III.1: Counts of SLPP graduates by year 197
III.2: Personal characteristics of sample APs as compared to other new APs.....198
III.3: School characteristics of sample APs as compared to other new APs 199
III.4: Characteristics of SLPP graduates by cohort.....200
III.5: Characteristics of APs by program 201
III.6: Characteristics of schools employing SLPP graduates.....202
III.7: PIIR Program Quality Scores 204
III.8: Coding for estimating selection PIIR score 205
III.9: Relationship between APs' perceptions of SLPP quality and PIIR scores.....206
III.10: Relationship between APs' evaluation ratings and PIIR scores 207
III.11: Relationship between APs' evaluation ratings and PIIR scores controlling for prior performance 208
III.A1: Coding for estimating curriculum PIIR scores 210
III.A2: Coding for estimating Partnerships PIIR scores 211
III.A3: Coding for estimating field experience PIIR scores 212
III.A4: Inter-rater reliability.....213
III.A5: Relationship between APs' perceptions of SLPP quality and PIIR scores.....214
III.A6: Relationship between APs' evaluation ratings and PIIR scores (first year) 215
III.A7: Relationship between APs' evaluation ratings and PIIR scores controlling for prior performance (first year) 216

INTRODUCTION

Assistant principals occupy a poorly defined space in the organizational structure of schools. They spend much less time working directly with students as compared to teachers, but they often spend more time dealing with the day-to-day operations of schools than principals. The education leadership and policy literature has largely overlooked the assistant principalship, partially due to the peculiar nature of the position. Moreover, the extant literature has primarily conceived of the assistant principalship as a stop on the way to the principalship. A fundamental issue that remains underexplored is whether the assistant principal is merely the *assistant to the principal* or the *assistant principal*. In other words, are assistant principals just executors of principals' visions or are assistant principals building leaders with their own contributions and unique functions in schools? This dissertation begins to address this gap in the literature by examining three aspects of the assistant principalship: assistant principals' contributions to school climate, assistant principals' time allocation to leadership duties, and assistant principals' preparation.

A growing research base explores the effect that schools have on student outcomes, yet the assistant principal is largely missing from these studies. The first chapter in this dissertation attempts to estimate the relationship between exposure to assistant principals and school climate outcomes. I examine the relationship between three different types of exposure to assistant principals (having an assistant principal, having one additional assistant principal, and having an assistant principal with a higher evaluation rating). I find that different types of assistant principal exposure have unique relationships with school outcomes. I find that having an assistant principal is related to reduced likelihoods of suspension and improved attendance. I also

find that having an assistant principal who has higher evaluation ratings is associated with lower likelihoods of teacher turnover and improved teachers' perceptions of school relationships. Finally, the improved student outcomes associated with having an assistant principal appear to be concentrated among Black students and elementary schools.

Although literature on assistant principals is generally thin, most of the quantitative studies are about the duties of assistant principals. This research largely suggests that assistant principals spend most of their time on discipline and administration. In the second chapter, I examine whether previously established patterns in assistant principals' duties hold in more recent years and in the context of Tennessee. I also explore whether assistant principals' time allocation is associated with several factors that are found to have relationships with teachers' and principals' time allocation. First, I find that assistant principals in Tennessee appear to allocate more time to instructional duties than any other type of duty. They also have substantial overlap in duties with principals, but they tend to allocate their time differently from principals. I also find that that the demographic characteristics of assistant principals and their schools are significant predictors of assistant principals' time on duties. Assistant principals in schools with higher achievement levels or principals with higher evaluation ratings allocate more time to instructional duties and less time to discipline. The results from this study also suggest that the time assistant principals allocate to duties is associated with the count of potential support staff and the time allocation of other formal leaders.

The differences between assistant principal and principal positions in the scope of their duties suggest school leadership preparation programs may want to give more thought to how they prepare their graduates for the assistant principalship. In the third chapter of this dissertation, I use a novel method for measuring the quality of school leadership preparation

programs that may be less intensive than the survey methods that provide the richest data on programs. I then use this measure to examine patterns in preparation program quality across Tennessee. I find that the novel measures of program quality differentiate programs along multiple features, and the measures of different program features generally covary with each other. The features that seem to be positively related to graduates' perceptions of their program quality are the curriculum and field experiences of programs, but these features have a negative relationship with evaluation ratings of their graduates when they are assistant principals. The quality of a programs' partnership with local districts is negatively related to graduates' perceptions of their programs, but their partnership quality is positively related to graduates' evaluation ratings as assistant principals even when controlling for past performance as teachers. This study suggests that the novel measure of program quality shows promise as a research tool for studying preparation programs. This study also suggests that the features that predict graduates' perceptions may not be the same as the ones that predict graduates' performance.

These three chapters suggest that the assistant principal has unique preparation needs, a distinct role, and a measurable contribution to schools. Although research on school leadership has centered on the principal, my dissertation suggests assistant principals have their own roles in and contributions to schools. This dissertation extends the literature on school leadership by providing evidence for the use of novel methods for investigating leadership preparation and more complex conceptions of what comprises the effect of school leadership that incorporate assistant principals.

CHAPTER I

Addition by Division or Quality over Quantity:

Estimating the contribution of assistant principals to school climate

School climate is a nebulous term used to capture a broad array of school environmental factors that students, teachers, and other school community members experience (Schweig et al., 2019). A growing body of literature suggests that school environments are critical to school outcomes, and a positive school climate is associated with improved student achievement, higher attendance rates, and fewer suspensions (Jonathan Cohen et al., 2009; Hoy et al., 1998; MacNeil et al., 2009; Uline & Tschannen-Moran, 2008). Research also suggests that teachers in schools with positive climates tend to be more committed, turn over less often, and engage in more collaboration (Collie et al., 2011, 2012; Preston & Barnes, 2017). Although a broad range of research has argued that school climate is important, school climate does not have a widely agreed upon definition (Jonathan Cohen et al., 2009; Schweig et al., 2019). Most scholars agree that school climate is multi-faceted, and the most widely cited definitions describe school climate as being comprised of four domains: academic, safety, relational, and structural (Jonathan Cohen et al., 2009; Thapa et al., 2013; M.-T. Wang & Degol, 2016). While few studies of school climate have focused on how the physical structures of the environment influence school outcomes (Duyar, 2010; Uline & Tschannen-Moran, 2008), most studies of school climate emphasize the psychological and behavioral aspects of school climate that impact school outcomes because those aspects tend to be more malleable (Bryk & Schneider, 2003; Collie et al., 2011; Hoy et al., 1998; Tschannen-Moran, 2014).

Research finds that school leaders play an especially important role in shaping those malleable aspects of school climate (Burkhauser, 2017; Griffith, 1999; Grissom et al., 2018; Liebowitz & Porter, 2019). Substantial evidence has accumulated to suggest that school leaders are central to setting the climate across multiple domains (Bryk et al., 2010; E. B. Goldring et al., 2015; Hoy & Tschannen-Moran, 2007; Rocque, 2010; Tschannen-Moran & Gareis, 2015). However, the available literature on school leaders' roles in managing school climate has centered on the principalship. The assistant principal (AP) has been largely overlooked in the school climate literature even though the roles and duties of APs are often closely related to school climate (Hausman et al., 2002; Sun, 2012).

According to the extant literature, the roles and duties most often assigned to APs are managing student discipline, supervising student activities, and handling administrative tasks (Austin & Brown, 1970; Glanz, 1994; Hausman et al., 2002; Koru, 1993). A more recent study of APs' time allocation found that teacher evaluation and instructional leadership are becoming essential components of their work as well (Sun, 2012). The execution of these various roles and duties performed by APs are intricately tied to the psychological and behavioral components of school climate. For example, the amount of exclusionary discipline used in schools is often used as an indicator of the climate of safety in a school (Astor et al., 2010; Bradshaw et al., 2014; Mitchell et al., 2018). Since APs often manage the student discipline in a school, they are likely to play an important role in making the school feel safe. APs have the potential to be an important contributor to school climate (Barnett et al., 2012; Marshall & Hooley, 2006), but there are no studies that suggest that APs have any relationship to positive school outcomes.

APs have the potential to improve school climate, but their ability to influence school climate is tied to how principals share leadership in schools. Several scholars lament how broad

principals' duties have become in recent years (S. H. Davis et al., 2005; Heck & Hallinger, 2009), and some suggest that sharing the responsibilities of leadership across multiple school personnel may be beneficial (Harris & Spillane, 2008; Mayrowetz, 2008; Murphy et al., 2009; Spillane, 2005; Zhu et al., 2018). The most prominent framework used in education leadership research to describe shared leadership is distributed leadership (Gronn, 2000; Heck & Hallinger, 2009; Spillane et al., 2001; Tian et al., 2016; Zuckerman et al., 2018). Some of the research on distributed leadership suggests that effective distribution of leadership activities in schools is associated, albeit indirectly, with improvements in student achievement (S. E. Anderson et al., 2009; Heck & Hallinger, 2009). Although many schools have APs, the literature on distributed leadership primarily examines the relationship between principals and teachers (Murphy et al., 2009; Spillane et al., 2001). Moreover, research on distributed leadership has generally examined leadership practices rather than potential organizational structures that may enable distribution of leadership tasks (Gronn & Hamilton, 2004; Spillane, 2005).

Having additional formal leaders in a school, like APs, may facilitate an organizational structure that enables more effective distribution of leadership. A handful of studies have examined how co-principalship may be a structure by which distributed leadership is "institutionalized" (Eckman, 2018; Gronn & Hamilton, 2004), but no studies I am aware of examine how the role of APs could potentially provide another structural mechanism by which distributed leadership is enacted. The studies of co-principalships highlight both the possibilities for organizational structures to support distributed leadership practices and how the structures do not guarantee distribution of leadership duties (Gronn & Hamilton, 2004). Although the AP role still maintains the traditional school hierarchy (Eckman, 2018), the AP's role as a positional

leader in the school creates a structure that should encourage principals to distribute their leadership.

Although APs could support distributed leadership practices and enhance school climate, research on APs is limited. This scarcity may in part be due to three critical challenges to estimating the relationship between exposure to APs and school. The first challenge is that it is difficult to disentangle the contribution of APs from the contributions of principals because research suggests that there is substantial overlap in APs' and principals' duties (Marshall & Hooley, 2006; Sun, 2012). Second, the indirect nature of APs' impacts on students means that a substantial amount of data are needed to estimate those impacts. Third, schools that have APs are fundamentally different from schools that do not have APs, so it may be difficult to partial out the influence of contextual factors from the influence of APs.

Although this study faces its own limitations in estimating APs' impacts on schools, the methods and data employed attempt to mitigate these three challenges. This study attempts to take a rigorous approach towards estimating the relationship between different measures of exposure to APs and school climate outcomes. To that end, it addresses four research questions. First, to what extent are student and teacher outcomes associated with having an AP? Second, to what extent are student and teacher outcomes associated with having additional APs beyond one? Third, to what extent are student and teacher outcomes associated with evaluation ratings of APs? Fourth, to what extent is the association between exposure to APs and student outcomes different for students from different demographic groups?

In the remaining sections of this chapter, I provide an overview of the relevant literature. Then I provide some context for how AP positions are allocated in the state of Tennessee. This section includes a discussion of how the factors associated with the allocation of APs may be

confounders in estimating the relationship between having an AP and school outcomes. I then turn to a description of the data and methods employed in this study. The methods section is followed by a description of the findings and a discussion of the conclusion I draw from those findings. Finally, I describe potential limitations of this analysis and conclude with future research that may address some of these concerns and extend this study.

I.1 Literature Review

I.1.1 School climate

Research on school climate does not provide a unified definition of the construct (Schoen & Teddlie, 2008; Van Houtte, 2005). What seems to be consistent across studies of school climate is that it is a multi-faceted construct that captures the *feeling* of a school environment (Jonathan Cohen et al., 2009; Thapa et al., 2013; M.-T. Wang & Degol, 2016). Cohen and colleagues (2009) describe school climate as the quality and character of school life. Freiberg and Stein (1999) define school climate as the "heart and soul of a school" (p, 11). Although there is some disagreement about the distinct domains of school climate, the domains as described by Cohen and colleagues (2009) are most commonly used when describing school climate. These domains are safety, teaching and learning, relationships, and environmental-structural. A key characteristic of school climate as described in the literature is that it is malleable (M.-T. Wang & Degol, 2016). This aspect of school climate is important because an underlying assumption of the many interventions aimed at improving school climate is that it can be shaped.

The malleability of school climate is a key difference between school climate and school culture (Schoen & Teddlie, 2008; Van Houtte, 2005). Although these terms are often used interchangeably, Van Houtte's (2005) comparison of these related constructs is especially helpful

in clarifying what differentiates climate from culture. Climate is described in the literature as *shared perceptions* while culture is defined as *shared beliefs and norms*. Climate is more susceptible to change; culture is more entrenched within an organization. Van Houtte (2005) argues that school climate is one piece of school culture, but Schoen and Teddlie (2008) argue that climate is a level of culture. In Schoen and Teddlie's conception of the relationship between climate and culture, climate reflects the underlying culture of a school. This perspective from Schoen and Teddlie (2008) is consistent with research that suggests school climate is made up of the shared perceptions about a school while school culture is made up of the shared norms and assumptions that inform those perceptions (M.-T. Wang & Degol, 2016).

Although school culture may be the more foundational construct, there is a mounting evidence base in education research to suggest that school climate is related to several important school outcomes (Thapa et al., 2013; M.-T. Wang & Degol, 2016). For example, research finds that a positive school climate is associated with reductions in student absence, improvements to student attendance, and reductions in suspensions (Lee et al., 2011; Purkey & Smith, 1983). Research also finds that positive school climate is related to improved outcomes for teachers (Gajda & Koliba, 2008; Hughes, 2012; Tschannen-Moran & Gareis, 2015). Specifically, a positive school climate is related to improved working relationships for teachers and decreased likelihoods of turnover (Burkhauser, 2017; Griffith, 2004; Grissom, 2011; Preston & Barnes, 2017).

Much of the research on school climate has emphasized the role of school leaders in shaping school climate (Bryk et al., 2010; Burkhauser, 2017; Hoy & Tschannen-Moran, 2007; Tschannen-Moran & Gareis, 2015). Research suggests that school leaders shape school climate by setting a vision for schools, evaluating teachers, cultivating relationships in a school, and

through strategic management of the school organization (E. Goldring et al., 2008; Grissom & Bartanen, 2019; Grissom & Loeb, 2011; Liebowitz & Porter, 2019). This research has exclusively examined the role that principals play in shaping school climate. Although some emphasis on the principal is appropriate because ultimate responsibility for school performance is vested in the principal, this conceptualization of school leadership envisions the activity of leadership as being constrained to one person. Although the influence of the principal should not be understated, research seems to have largely ignored the role that other formal school leaders, like APs, can play in shaping school climate.

I.1.2 Distributed leadership

One possible mechanism by which APs can improve school climate is by sharing in the leadership tasks and responsibilities of the principal (Gronn, 2000; Mayrowetz, 2008; Spillane, 2005). Although there are many theories about how leadership is shared across multiple personnel in an organization, distributed leadership is the theory with one of the largest research bases in education leadership research (Harris, 2013; Tian et al., 2016; Zuckerman et al., 2018). Tian and colleagues (2016) argue that the literature on distributed leadership can be generally categorized into two paradigms: descriptive-analytic and prescriptive-normative. Under the descriptive-analytic paradigm, researchers attempt to explain the practice of leadership using a distributed perspective (Gronn, 2002; Spillane et al., 2001). Distributed leadership theory suggests that leadership activity can be described as the interaction of leaders, followers, and situations (Spillane, 2005). Under the prescriptive-normative paradigm, researchers evaluate the effects of specific leadership practices that they describe as distributed leadership and advocate for the implementation of those leadership practices (Harris, 2009; Heck & Hallinger, 2009). In

this prescriptive of distributed leadership, leadership will be more effective if it is structured to be distributed widely across school personnel (Harris, 2009; Murphy et al., 2009).

Several academic studies have used distributed leadership as a framework, but few studies have actually examined the relationship between school structures that facilitate distributed leadership and school outcomes using quantitative methods (S. E. Anderson et al., 2009; Heck & Hallinger, 2009; Hulpia et al., 2010; Tian et al., 2016). Hulpia and colleagues, find that distributed leadership is related to improved teacher commitment and job satisfaction in Belgian secondary schools. Anderson and colleagues (2009) find that distributed leadership has a weak relationship with student achievement and that the correlations vary in direction across school subjects in five British schools. The only large-scale study of distributed leadership was conducted by Heck and Hallinger (2009), and it finds evidence to suggest that distributed leadership has a significant but indirect relationship with student achievement. Although there are not many studies that suggest there is a positive relationship between distributed leadership and school outcomes, the studies that exist are fairly convincing and provide an intuitive explanation for how principals influence school outcomes. There are more qualitative and theoretical studies that suggest distributed leadership could make school leadership more effective (Gronn & Hamilton, 2004; Grubb & Flessa, 2006; Murphy et al., 2009; Smylie et al., 2007), but more quantitative research is needed to evaluate under what conditions distributed leadership could be most effective in improving school environments. What is especially needed are studies of different organizational structures that vest leadership responsibilities in multiple formal leaders and how those structures influence school performance (Grubb & Flessa, 2006). Most of the research on distributed leadership focuses on how leadership is distributed across

principals and teachers (Tian et al., 2016); no studies are available that examine how leadership is distributed between principals and APs.

Perhaps it is assumed that principals distribute leadership with APs, because definitionally APs are meant to assist principals in their work. However, much of the research literature on school leadership implies that any leadership effects on schools are attributable to principals (Burkhauser, 2017; Heck & Hallinger, 2009; Sorensen et al., 2020). These two conceptions of leadership seem to be at odds with one another. If leadership is distributed between principals and APs, then some of the influence of school leadership on school outcomes should be attributable to APs. If the effect of school leadership is solely attributable to the principal, then there must not be any substantive relationship between the activities of APs and school outcomes. While the evidence in this study cannot explain the proportion of variation in school outcomes that is attributable to APs and principals, it can test whether there is any observable change in teacher and student outcomes associated with changes in exposure to APs. If APs do not matter to school leadership, then there should be no observable differences in teacher and student outcomes associated with changes in exposure to APs. Although observable differences in teacher and student outcomes associated with changes in exposure to APs would not prove conclusively that APs matter for schools, these changes would suggest that there is potential to estimate a causal impact of APs on schools.

I.1.3 Research on APs

Research on APs has examined a range of topics including their work duties, preparation, socialization, and paths to the principalship (Allen & Weaver, 2014; Busch et al., 2012; Fuller et al., 2016; Gates et al., 2004; Marshall & Hooley, 2006; Sun & Shoho, 2017). Unfortunately, the

empirical literature on APs does not provide much evidence as to how APs influence distributed leadership, school climate, or school outcomes. The most robust areas of existing research on APs are on their preparation, socialization, and duties (Oleszewski et al., 2012). Of these strands of research, studies of APs duties have the most relevance to the research questions explored in this chapter. The research on APs' duties explores, primarily through surveys of APs, what they do in their jobs (Glanz, 1994; Hausman et al., 2002; Koru, 1993; Sun, 2012; Sun & Shoho, 2017). Many of these studies suggests that principals are the single most important factor in determining the span of APs' duties (Austin & Brown Jr, 1970; Glanz, 1994; Koru, 1993). Researchers argue that principals assign APs duties they do not want or are not essential to instruction (Glanz, 1994; Koru, 1993; Marshall & Hooley, 2006).The early research in this area finds that APs were asked to perform a number of duties, but these duties were most often related to managing discipline issues among students and performing administrative duties like making the master teaching schedule (Austin & Brown Jr, 1970; Glanz, 1994). Part of the problem with understanding the AP position is that it is generally ill-defined (Watson, 2005). More recent studies of APs' duties find that APs are now expected to perform more instructional leadership duties like observing teachers and coaching them on their instructional practice (Hausman et al., 2002; Petrides et al., 2014; Sun, 2012). These aforementioned studies are largely limited to simple descriptions of APs' duties, and few studies extend their analysis beyond simple descriptions.

The existing research on APs' duties may not explicitly explain how APs contribute to school outcomes, but it can point to potential pathways through which APs influence schools. Research suggests that APs spend their time managing discipline, attending to administrative duties like attendance, and engaging in instructional leadership tasks like observing instruction

(Glanz, 1994; Hausman et al., 2002; Koru, 1993; Sun, 2012), and these tasks all have a clear connection to school climate. For example, a dimension of school climate is safety, and the effective implementation of school discipline shapes how safe students and teachers feel in the school. Since one of the main duties of APs is school discipline, APs should have an impact on school safety and by extension school climate (Hausman et al., 2002; Sun, 2012).

Prior research does not directly examine the relationship between exposure to APs and school safety, but several studies find that principals' characteristics are related to the likelihood students experience exclusionary discipline (Findlay, 2015; Kinsler, 2011; Sorensen et al., 2020). One recent working paper finds that a principal's underlying propensity to use exclusionary discipline practices is related to an overall reduction in the number of infractions in the school, but it is also related to lower graduation rates, more juvenile justice complaints, and lower attendance and test scores for students that commit minor infractions. If a principal is more likely to exhibit bias when assigning discipline to Black and Latinx students, there are additional negative consequences for those students. Interestingly, the authors do not find any differences in the relationship between principals' propensity to use exclusionary discipline and students' likelihood of experiencing exclusionary discipline across schools with different numbers of APs. The authors argue that APs do not influence the student discipline process because they do not find any differences in the relationship between schools with different numbers of APs, but the authors do not actually demonstrate that the number of APs does not have any relationship with the likelihood a student experiences exclusionary discipline.

Qualitative studies suggest APs can impact outcomes for students, and research should explore if there are observable differences in outcomes for students related to their exposure to APs. For example, some qualitative studies suggest that APs may be especially important for

students who have been traditionally marginalized (Carpenter et al., 2017a; Clayton & Goodwin, 2015; Moore, 2013). Carpenter and colleagues (2017) find evidence to suggest that APs can play an especially important role in low-achieving schools by serving as community builders. Moore (2013) finds evidence to suggest that some APs of color use their roles to address educational inequities in schools. Clayton and Goodwin (2015) find in two schools that white APs can also be effective in serving the needs of a racially diverse student population if they exhibit cultural competences in their implementation of discipline. These studies argue that APs have the potential to have bigger impacts on students of color and students that have been historically marginalized. These studies also suggest that not all APs will have the same type of impacts and some APs may be more effective than others. In the aforementioned studies, the characteristics that differentiate APs from each other are cultural competence and race, but it may be possible that other characteristics of APs are associated with student outcomes.

The literature on APs' duties suggest that their span of work touches many aspects of school climate, but no quantitative research to date has examined the relationship between APs activities and school outcomes. Research on the activities of principals in school discipline, a primary duty for APs, suggests that principals have a measurable relationship with the outcomes associated with that domain (Kinsler, 2011; Skiba et al., 2014; Sorensen et al., 2020). These studies largely define school discipline decisions as the primary responsibility of principals. If school discipline decisions are primarily the responsibility of principals, then why do APs consistently report across studies that discipline is one of their primary duties? In terms of ultimate responsibility, principals are responsible for how discipline is applied in their schools, but the assumption that principals are making the bulk of the discipline decisions does not comport with the research on APs' duties. Moreover, qualitative research suggests that APs may

be especially influential in implementing discipline for students of color who often experience disproportionately more incidents of exclusionary discipline (Carpenter et al., 2017a; Moore, 2013; Williams et al., 2020). This prior research suggests that examining disciplinary outcomes and potential differential relationships for students of color and other student groups that are often disproportionately impacted by exclusionary discipline are important (Skiba et al., 2014). So, I explore these issues by examining the likelihood of receiving any in-school or out-of-school suspensions associated with changes in AP exposure and conducting sub-group analyses by race, FRPL, and IEP status for those outcomes.

I.1.4 School climate, distributed leadership, and the role of the AP

The available research on APs, distributed leadership, and school climate help to identify a theoretical pathway by which APs may influence school outcomes. The gaps in the research also point to areas that need further investigation. The research on school climate suggests that school environments are critical to positive student and teacher outcomes (Burkhauser, 2017; Thapa et al., 2013; M.-T. Wang & Degol, 2016), but this research only examines the role of principals in shaping school climate. The distributed leadership research focuses on what principals do as gate keepers in distributing leadership (Bush & Glover, 2012; Tian et al., 2016), but it does not examine how organizational structures, like having an AP, facilitate distribution of leadership responsibilities. The research on APs is generally thin, and it does not address the efficacy of APs (Oleszewski et al., 2012; Sun & Shoho, 2017). The extant literature on APs' duties does not examine the relationship between exposure to APs and outcomes of teachers and students.

This study attempts to bridge these areas of research by estimating the relationship between exposure to APs and school climate outcomes. This study attempts to identify the

contribution of APs to school climate, and it points to potential pathways for APs to impact schools. There is little existing research to guide this analysis; however, utilizing the relevant literature on school climate, distributed leadership, and AP duties I develop a potential framework for conceptualizing how APs contribute to school outcomes. This conceptual framework is presented graphically in Figure I.1.

First, this diagram conceptualizes the work of APs as part of the larger activity of the school leadership team as suggested by the literature on distributed leadership and co-principalships (Gronn, 2002; Spillane, 2005). It also theorizes that part of the influence of principals on schools goes through APs. In this diagram the pathway for APs to influence student and teacher outcomes is through their duties. The research on APs' duties suggest that their duties are in domains closely related to several facets of school climate (Glanz, 1994; Hausman et al., 2002; Koru, 1993; Marshall & Hooley, 2006; Sun, 2012; Williams et al., 2020). Though there may be many characteristics of APs who may moderate the effectiveness of their leadership activities, one potentially important characteristic is their effectiveness. The leadership activities of APs then influence school climate, and school climate affects teachers and students.

The AP's impact on teachers and students can be measured by the outcomes of student attendance, suspensions, teachers' perceptions of school relationships, and turnover. The model also acknowledges that the relationship between AP work and student outcomes may be direct rather than indirect through school climate. This conceptual framework is constructed based on prior research in a number of areas related to school leadership, but this model is only one of many possible models. This study is not able to uncover the central mechanisms through which APs influence school climate and student outcomes. It attempts to estimate a relationship between the initial input of exposure to APs and the output of students' and teachers' outcomes.

If an observable relationship between exposure to APs and outcomes related to school climate can be established, it would lend credence to this type of conceptual model.

I.2 APs and the Tennessee context

A central challenge for this study is that APs are not randomly distributed across schools. There are factors that are likely to be related to both the assignment of APs and the school outcomes examined in this study. Some common factors associated with the allocation of personnel include school enrollments, grade levels, locales, and the proportion of students needing additional supports (Kelly & Chesser, 2019). Each of these factors are likely to influence both the exposure of APs to schools and school outcomes. For example, in many policy contexts a school's enrollment is positively correlated with allocation of an AP, and prior studies find a negative relationship between enrollment size and school climate (Kelly & Chesser, 2019; Thapa et al., 2013). Simply due to the correlation that allocation of APs and school climate have with enrollment, allocation of APs and school climate will have a negative correlation.

In Tennessee, the state government provides districts with funding based on personnel allocations for schools (Tennessee State Board of Education, 2019). The state has a formula for personnel allocations based on the average daily membership in the school (Tennessee Department of Education (TDOE), 2014). Average daily membership is a measure of school enrollment that weights the number of students enrolled in the school by their average attendance (TDOE, 2014). Appendix Figure I.A1 presents the formula for determining how many APs the state funds based on average daily membership in elementary schools and secondary schools. Generally, elementary schools in Tennessee serve grades PK to 5 and secondary schools in Tennessee serve grades 5 to 12. The grade level classification of schools does not strictly adhere

to these divisions. The state funds instructional personnel, which its formula considers APs to be, at 75 percent of the state's estimated cost for personnel, and expects school districts to match state spending to fund the remaining 25 percent of the costs. Although the state's formula is built on a district match to state spending, the state provides additional funds based on an equalization formula to districts that do not have the resources to fund the match. Districts are allowed to raise and spend additional funds for schools, but the state provides the minimum funds for staff allocations through the Basic Education Program formulas (Polanchek, 2019).

Several other states use similar formulas to Tennessee's for determining how to allocate APs. Kelly and Chesser (2019) provide an overview of how administrators are allocated across multiple states. In their analysis, Kelly and Chesser (2019) find that several states fund school administrators based on enrollment. New Jersey provides per student funding with some state-level guidance on how districts should allocate funds for APs through an illustrative example. North Carolina funds months of employment for APs based on enrollment. Only Virginia has state level mandates that require minimum numbers of APs to be assigned to schools based on enrollment. Across all the states surveyed in their analysis, districts have considerable control in determining how to allocate APs (Kelly & Chesser, 2019).

Tennessee's funding formula is not a spending formula (Tennessee State Board of Education, 2019). The state provides funds for personnel according to the Basic Education Program, but it allows districts to determine how they want to distribute those funds to schools. There is some variation across districts in how they distribute funds to schools. As an example, I include in Appendix Figures I.A2-4 the formulas that three of the largest school districts in Tennessee use for distributing funds to their schools. These appendices illustrate how a district may diverge from the state's funding formula. Knox County's budget allocates teachers in

schools with high proportions of "at-risk" students at a lower student to teacher ratio than schools with lower proportions of "at-risk" students. In both Metro Nashville Public Schools and Shelby County Schools, funding is allocated using a student-based budgeting model. As the funding formulas for these districts illustrate, districts are not expected to strictly adhere to the state formulas for allocating APs.

Table I.1 presents a cross tab of schools with zero or one AP and whether or not schools qualify for AP FTEs using the state's Basic Education Plan formula. This table shows that about 20 percent of schools have an AP if they are funded for any FTEs[1] of AP funding and 35 percent of schools have an AP even if the state formula does not provide funds for an AP in that school. Table I.1 also shows that if being funded for an AP position is defined as having at least 1.0 FTE for an AP, only 5.7 percent of schools would be both funded for an AP position and have an AP, and 49.5 percent of all schools would not be funded for an AP position and have an AP. Table I.1 suggests that more schools have an AP even though the state does not provide the funds for an AP in the school. Districts apparently do not closely follow the state's Basic Education Plan formula when making AP staffing decisions. Despite what appears to be substantial "non-compliance" to the Basic Education Plan by school districts, the number of APs in a school has a relatively strong correlation with student enrollment at .79 ($p < .001$).

One potentially compelling approach to estimating the relationship between having an AP and outcomes is to instrument for the number of APs in a school with the number of AP FTEs funded under Tennessee's BEP. This approach should in theory leverage the variation in the number of APs that is exogenous to student and teacher outcomes. When I attempt to

[1] FTE refers to full-time equivalents and is used to describe the type of employment a person has. 1.0 FTEs is the equivalent of 1 full time employee in that position.

estimate instrumental variables models using the BEP AP FTEs as an instrument for the number of APs, I find that the number of funded AP FTEs is a weak predictor for the number of APs in a school. The cross tabulation of the number of APs in schools and funding for AP positions under the BEP reported in Table I.1 points to a possible explanation for the weak relationship between the BEP formula and the actual number of APs in a school. As mentioned above, few schools seem to actually "comply" with the assignment of APs according to the BEP. The results from the instrumental variables estimation and the first stage F-statistics for these models are presented in Appendix Table I.A1. All of these models have first stage F-statistics below the traditional threshold of 10 to avoid finite sample bias. As suggested by the descriptive results reported in Table I.1, it seems that the BEP funding is only part of a complex process by which districts assign APs to schools. An example of the type of unobservable factors that districts may consider in assigning APs is a district's approach to succession planning. If a district is anticipating a principal retirement in upcoming years, an AP may be assigned to a school that is not funded for an AP FTE under the BEP to prepare a successor the principal. These unobservable processes seem to explain substantial variation in the number of APs in a school beyond variation that can be explained by the funded number of APs.

Since in Tennessee the likelihood of exposure to APs is related to school enrollment, and research suggests school climate is correlated with a school's enrollment (Thapa et al., 2013), the relationship between exposure to APs and school climate will be biased without properly accounting for enrollment of a school in estimation. The estimation approach I employ includes several of the variables that might be related to both allocation of APs and school climate including average daily membership. However, it may be that districts systematically allocate APs to schools in anticipation of how changes to schools also impact school climate. For

example, a school district could systematically assign APs to schools that are experiencing a negative trend in school climate. The estimation strategy I use would attribute the worsening school climate to the assigning of an AP, even if the changes to school climate are a function of a pre-existing trend in climate. Since I am not able to control for the exact selection criteria for the allocation of APs, the results from my analysis should be considered descriptive.

To examine the extent to which pre-existing trends may bias the results in this study, I have conducted two analyses. For these analyses I identify a smaller sub-sample of schools that experience a change in having an AP but do not experience any changes in APs for at least three years prior to a change. I separate schools that gained an AP from schools that lost an AP. For this smaller sample of schools, I can examine descriptively whether schools that gain or lose an AP have trends in their outcomes in pre-change periods that would suggest that pre-existing trends may be related to changes in the number of APs. Moreover, I can formally test whether there are significant deviations in pre-change periods as compared to the year prior to the change in having an AP. I describe these results in the section on sensitivity analyses following the main findings of this study.

I.3 Data and Measures

The data for this study comes from the administrative data provided by the Tennessee Department of Education (TDOE) through the Tennessee Education Research Alliance (TERA) at Vanderbilt University. This study focuses on the school years ending in 2012 to 2019 because the variables needed for this study are consistently available for the entire time frame. The first set of data are found in the staff files that contain the demographic characteristics (gender, race, age, degree attainment, and experience as an educator) and school linkages for all public school

staff in Tennessee from 2002 to 2019. From the data on positions, I construct the job histories of all staff and create measures of experience as a teacher, as an AP, and as a principal in Tennessee. The position data are also used to create school-level variables for the average characteristics of teachers and APs in the school.

The administrative data from TDOE also contains rich information on students in Tennessee's public schools including students' demographic characteristics (gender, race, FRPL status, and having an IEP[2]), disciplinary outcomes, attendance and linkages to teachers and schools. I use these student data in the student analyses and to create school-level measures of the average daily membership and the average demographic characteristics of students.

Additionally, evaluation data on school staff are available starting from 2012 through the Tennessee Educators Acceleration Model (TEAM) system of evaluations. There are separate evaluation formulas and rubrics for teachers and school leaders. School leaders in Tennessee are evaluated annually on school-wide achievement and subjective ratings given by designated supervisors; in the case of APs this rater is most often the school's principal. For this study, I focus on the subjective ratings of APs as a measure of AP effectiveness. The TEAM rubrics for school leaders ask APs' supervisors to rate them on a five-point scale across multiple dimensions related to the leadership practices of APs. Prior research on the TEAM ratings for school leaders suggest that scores are both internally consistent and stable over time (Grissom et al., 2018). Since Grissom and colleagues (2018) find evidence to suggest that the different domains on the TEAM rubric for school leaders derive from one underlying factor, I use the average ratings across all the domains. I keep the scores in their original one to five rating scale for a more intuitive interpretation of the results.

[2] In the analyses for this dissertation, I define having an IEP as a student who has an individualized education plan for special education services but does not receive gifted services.

The last data source used in this study is the survey data available from 2012-2019. The variables that I create from the survey data are measures of school climate as reported by teachers. These measures are especially important because school climate is defined as the perceptions of the school environment. From 2012-2014, TDOE administered a survey of all teachers and administrators in Tennessee as part of the First to the Top initiative. These surveys were aimed primarily at capturing teachers' and administrators' perceptions of the TEAM evaluations, but measures of school climate were also collected each year. Starting in 2015, TERA administered a revised survey called the Tennessee Educator Survey (TES) that captured a broader range of perceptions related to the work of teachers and administrators. The TES is revised every year, but each administration of the TES includes some questions about school climate. The response rates for the surveys and the survey questions included in each year of the Tennessee surveys are listed in Appendix Table I.A2. For simplicity I refer to all of the surveys as the TES throughout the remainder of this chapter.

I.2.1 Measures of exposure to APs

There are three different types of exposure to APs that I examine in this study. The first type of exposure is switching between having no APs and having one AP. This indicator takes the value of zero for years when students and teachers have no APs and a value of one for years when they have one AP. I exclude the observations of schools in years when they have more than one AP. Table I.2 presents descriptive characteristics of schools in years with no APs as compared to schools in years with one AP. This comparison highlights the differences between schools with APs and schools without APs. Schools with APs tend to serve higher grade levels, are in suburban neighborhoods, and are larger than schools without APs. Although the comparison in

Table I.2 is illuminating, the estimation strategy employed in this analysis does not leverage all the data in estimating the relationship between changes in exposure to APs and changes to school outcomes. Schools that never experience a change in their exposure to APs are not leveraged in the estimate of exposure to APs. Therefore, Table I.3 presents a comparison of the effective sample that is leveraged for estimation as compared to the total study sample, and Table I.3 shows that the effective sample is comprised of more elementary schools, more urban schools, and smaller schools. Lastly, Table I.4 compares the characteristics of schools in the effective sample that gain an AP to schools that lose an AP. There are no measurable differences between the types of schools that experience these different changes in exposure to APs.

The second type of exposure to APs is having multiple APs. The measure I use in this study is the count of APs in a school.[3] Table I.5 presents the descriptive characteristics of schools with one AP as compared to schools with multiple APs. Schools with multiple APs tend to serve upper grades, are more likely to be suburban or urban, and are larger. The final type of exposure to APs is exposure to APs of different effectiveness ratings. The reason I include an analysis of this type of measure is because the relationship between having an AP and school outcomes may depend in part on the effectiveness of the AP. I use the average TEAM evaluation ratings because APs are supposed to receive multiple evaluation ratings and the average rating may be closer to the true effectiveness of APs than a single evaluation rating. I only compare schools with one AP in this analysis so as to differentiate between changes in the number of APs from changes in the quality of APs. Table I.6 provides a descriptive comparison of schools with APs who have low evaluation ratings from schools with APs who have high evaluation ratings. For

[3] Although I run all the same analyses with the ratio of teachers to leaders in the school and students to leaders in the school, the patterns are clearest when using the simple measures of the count of APs.

this table, I categorize APs with low evaluation ratings as APs with average scores below four and APs with high evaluation rating as APs with average scores of four or higher. This table suggests that APs with high evaluation ratings are more likely to work in suburban schools with fewer students of color, fewer students that are FRPL eligible, and fewer students with IEPs.

I.2.2 Outcome Measures

In this study, I examine the relationship between exposure to APs and student and teacher outcomes that are associated with changes to school climate. The four types of outcomes are student attendance, student suspensions, teacher turnover, and teachers' perceptions of school relationships. I use these four types of outcomes because research suggests that they are connected to improved school leadership and improvements to school climate (Heck & Hallinger, 2009; Thapa et al., 2013; Tian et al., 2016; M.-T. Wang & Degol, 2016), and these four outcomes are likely to have a relationship with the specific duties that APs spend the most time on (Hausman et al., 2002; Oleszewski et al., 2012; Sun, 2012). The results of my analysis of APs' time allocation in Tennessee presented in Chapter II suggests that APs in Tennessee report spending the most time on similar tasks as has been identified in prior research. APs in Tennessee report spending the most time on student discipline, administrative duties, and instructional leadership.

For student outcomes, I examine two different measures of attendance (absence rates and chronic absenteeism) and two different measures of suspensions (in-school and out-of-school). I focus on attendance rate and chronic absenteeism because Bartanen (2020) finds evidence to suggest that principals have a measurable impact on these measures of attendance. If principals have an impact on attendance it is more likely that APs also impact attendance. The specific

measures of attendance rate and chronic absenteeism are especially important under ESSA. Being chronically absent is defined as having 18 or more absences in a school year. Research on school discipline finds that leaders are critical in the use of suspensions (Rocha & Hawes, 2009; Rocque, 2010; Sorensen et al., 2020). I specifically examine suspensions in this analysis because they are most widely used form of exclusionary discipline and more likely to have enough instances to detect differences associated with changes in exposure to APs.

I also analyze two teacher outcomes: teacher turnover and teachers' perceptions of school relationships. Teacher turnover is operationalized as the probability of not being employed at the same school in the following year. This measure does not differentiate between changing schools or leaving the state public school system. Based on the framework of the study, there is no reason to believe that APs would have a specific impact on a particular type of turnover behavior other than leaving the school.

The measure of perceptions of school relationships is obtained from teachers' responses on the TES. Appendix Table I.A2 lists all the questions used in creating this measure. These questions ask teachers to describe how students and adults relate to one another. Since the questions are not consistent across years, they may be capturing different constructs. To check that the measures of relationships are related to one another across years, I conduct a simple analysis that predicts a school's perception of relationships using the prior year's perception of relationships. This analysis suggests that there is a .38 correlation year to year in teachers' perceptions of school relationships that is significant at the .01 level even though the questions change. Within each year, the questions have an average inter-item correlation of .38 and an average Alpha coefficient of .95. The estimates of the reliability of the relationship measure for each year are presented in Appendix Table I.A3. I also use principal component analysis (PCA)

to determine if all the individual survey items are measuring one underlying construct. The results from these PCA analyses are presented in Appendix Figure I.A5, which plots the eigenvalues from the PCA on the factor number. Each year, there is only one factor with an eigenvalue above one. I obtain the estimated factor scores for each teacher from this PCA and standardize those scores, so that they have a mean of zero and a standard deviation of one.

Table I.7 reports the mean outcomes and standard deviations for teachers and students in each of the analytic samples. This table suggests that attendance rate for students is relatively high at just below 95 percent. Consistent with Bartanen's (2020) findings, I find that students in upper grades are more likely to have lower attendance rates and are more likely to be chronically absent. A similar pattern is true of suspensions. Students in upper grades are more likely to have suspensions and are suspended more frequently than students in lower grades. On average, teacher turnover is approximately 16 percent annually across the study samples, which is similar to prior research on teacher turnover in Tennessee schools (Grissom & Bartanen, 2019; Springer et al., 2016). Teacher perceptions of relationships in schools in the analytic sample has a similar distribution to the total sample, with a mean of about zero and a standard deviation close to one. These patterns in suspensions and attendance raise the concern that it may be that any of the significant associations between AP exposure and outcomes are concentrated in particular grade level. I assess the extent to which patterns differ across grade levels by conducting subgroup analyses by grade level.

I.4 Methods

To estimate the variation in student and teacher outcomes attributable to exposure to APs, I employ a person-by-school fixed effects estimator. Formally, I estimate a model of the following form:

$$y_{ist} = \beta_0 + \beta_1 HasAP_{st} + X'_{ist}\Gamma + \mu_{is} + \tau_t + \epsilon_{ist} \quad (1)$$

Where y_{ist} is an outcome[4] for person (student or teacher) *i* in school *s* and year *t*, $HasAP_{st}$ is an indicator for having an AP, X'_{ist} is a vector of controls including student characteristics, teacher characteristics, and principal characteristics, μ_{is} represents a vector of person-by-school fixed effects, τ_t is a vector of year fixed effects, and ϵ_{st} is an idiosyncratic error term. Although all the included covariates have the potential to bias results if they are excluded from the model, a particularly important variable in the model is principal turnover in the prior year. This control is important because the estimate on AP exposure is capturing part of the effect of leadership on school climate. If changes to principals are systematically more likely to occur concurrently, the estimates on having an AP will pick up the effect of principal turnover if it is not controlled for in the model. Therefore, I include a control for principal turnover in the year prior, which captures a change in principal between the prior school year and the current school year. The model employed in this study leverages variation within person while in the same school. The coefficients from these models should be interpreted as the marginal differences in individuals' outcomes as compared to themselves in years when they were in the same school. The reason person-by-school fixed effects are important to this model is because the most likely scenario by

[4] The estimates for models that have turnover, any ISS, any OSS, and likelihood of being chronically absent as the outcome impose the assumption that the relationship between the outcome and regressors is linear. These linear probability models are used for ease of interpretation, but as a check I also conducted logistic regressions to account for the non-linearity in the outcome. The results are substantively similar to the results presented in the tables in this chapter.

which students and teachers will experience changes in exposure to APs is due to moving schools. The strength of the person-by-school fixed effects model is that it avoids conflating changing schools and changes in exposure to APs.

The standard errors are estimated to be robust to heteroskedasticity and clustering at the fixed effect unit. Although unit level fixed effects are included in this model, simulation studies have suggested that a correlation in error terms can be induced with the inclusion of cluster unit fixed effects and time fixed effects. The cluster and heteroskedasticity robust standard error estimates produced by Stata have been found to be robust to this induced correlation (Abadie et al., 2017). There is some debate in the applied econometrics literature over the inclusion of fixed effects and cluster robust standard error estimates, but most scholars seem to agree that clustering standard errors allows for more conservative inferences that are less likely to suffer from Type I error. There are also no definitive rules about the level at which standard errors should be clustered.

Cameron and Miller (2015) argue that the level at which standard errors should be clustered boils down to the tradeoff between bias and variability. The higher the level at which standard errors are clustered, the less likely the estimates are to suffer from bias but the standard error estimates will have more variability. In general, the authors suggest clustering at higher levels when analyzing nested data. They also suggest for the cases of models that include unit specific fixed effects that standard errors should be clustered at the unit of the fixed effect. In theory, the unit specific fixed effects should account for any common shocks across observations within the same cluster unit. The literature does not seem to provide a definitive answer for what level standard errors should be clustered over when using unit specific fixed effects and analyzing nested data. A recent working paper suggests that standard errors should be clustered

by geographic regions when analyzing state level interventions (Powell, 2017). However, this working paper addresses the specific case of analyzing state panel data that are characterized by a small number of clusters and potentially large number of units within each cluster. Other studies suggest that the procedures for accounting for clustering should be different when there are a large number of clusters and a small number of units within each cluster as is the case in this study (Abadie et al., 2017; Cameron & Miller, 2015). Although the main results make inferences based on standard errors clustered at the fixed effect unit (person-by-school), I also estimate the standard errors at the school level as a robustness check.

To answer the second research question, I estimate the coefficient on having one additional AP in a school. These models are of the form in Equation 2.

$$y_{ist} = \beta_0 + \beta_1 APCount_{st} + X'_{ist}\Gamma + \mu_{is} + \tau_t + \epsilon_{ist} \quad (2)$$

All terms in Equation 2 are the same as in Equation 1, except the treatment indicators in the previous equation have been replaced by $APCount_{st}$ which is a measure of the number of APs in a school. The estimates of β_1 capture the marginal difference in a person's outcomes associated with having one additional AP when compared to themselves while in the same school.

One potential concern with the analyses for addressing the first two research questions is that AP quality may also be related to the outcome measures. My third research question attempts to address this concern by estimating the relationship between AP quality and the measures of student and teacher outcomes. This analysis should reveal whether students and teachers have better outcomes when they have more effective APs. The measure of AP quality that I utilize in these analyses is the evaluation rating of APs. To answer the third research question of this study, I estimate models of the following form:

$$y_{ist} = \beta_0 + \beta_1 APQuality_{st} + X'_{ist}\Gamma + \mu_{is} + \tau_t + \epsilon_{ist} \quad (3)$$

The *HasAP* indicator in Model 1 is replaced by the measure of $APQuality$ but all other terms are the same as in Equations 1 and 2. As a robustness check I use the prior year's evaluation rating of the AP in place of the same year's evaluation of the AP and find substantively similar results. I restrict the sample in these analyses to schools with only one AP because I want to minimize the potential for confounders.

The final analysis examines if there are any differences in relationships between exposure to APs and student outcomes for students from historically underserved identities. For these analyses, I run equations 1-3 on the sub-sample of students that are white, Black, Latinx, other students of color, FRPL eligible, and have an IEP. This analysis should identify sub-samples for which having an AP has differential relationships with the student outcomes.

I.5 Findings

The first set of results address the first research question, which explores the relationship between teacher and student outcomes and having an AP. Table I.8 presents the results from four different models to demonstrate how the estimates on the coefficient of interest change as different fixed effects are included in the model. The models in Table I.8 estimate the relationship between having an AP and teachers' perceptions of their school community. Column 1 shows the coefficient when only covariates are included in the model. Based on the results from column 1, it appears that having an AP is associated with having .0545 SDs lower perceptions of school climate compared to not having an AP. However, Table I.2 suggests that schools that have an AP are different from schools that do not have an AP and these observed differences suggest that there may also be unobserved differences between schools that have an AP and schools that do not have an AP. Column 2 includes school fixed effects and the

coefficient has reduced in absolute magnitude. Although the coefficient in column 3 is similar in magnitude to the coefficient in column 2, the standard errors have become much larger. This increase in the standard errors is expected considering that coefficients are now estimated on the variation within person. Also, the Adjusted R-Squared jumps to .52 in column 3 from .19 in column 2, suggesting that there is important variation in perceptions of climate among teachers that is not captured by the school fixed effects. Column 4 shows the coefficient when person-by-school fixed effects are included and there is a slight decrease in the absolute magnitude of the coefficient and an increase in the size of the standard error. This change suggests that some of the within person variation associated with having an AP can be explained by differences associated with being in different schools. Although the differences between columns 3 and 4 are not dramatic, they suggest that there is a difference in estimates and inferences made using the two approaches. Although the generalizability of the estimates suffers by focusing on the within person by school variation, the estimates from column 4 are the least likely to suffer from bias from contemporaneous changes that occur simultaneously with changes in exposure to APs.

Overall, Table I.8 suggests that teachers who have an AP have significantly less positive perceptions of their school community than teachers in the same school when they do not have an AP. However, the same teacher in the same school has no measurable differences in their perceptions of the relationships in their school community when they have an AP as compared to when they do not have an AP. Although the coefficient magnitudes are similar across models 2-4, the results from column 4 suggest we cannot infer that difference in perceptions associated with having an AP within teachers while in the same school does not occur by chance. The remainder of the results only show the person-by-school fixed effects estimates, so all results are relative to the same person while in the same school.

Table I.9 reports the estimated association between having an AP and school climate outcomes. The models in Table I.9 restrict the sample to schools that either have zero or one AP. Column 1 as reported in Table I.8 shows that there are no measurable differences for teachers' perceptions of school relationships in years when they have an AP as compared to years when they do not have an AP. There are also no measurable differences in having an AP and the likelihood of turnover in the following year. Having an AP is not related to a measurable difference in students' likelihood of out-of-school suspensions, but having an AP is associated with a lower likelihood of receiving any in-school suspensions, a lower likelihood of being chronically absent, and higher attendance rates.

The results that address the second research question in this study are reported in Table I.10. Table I.10 reports the relationships between teacher and student outcomes linked to school climate and having more APs beyond one. The sample for these models is restricted to schools in years when they have one or more APs. The coefficients in this table are the estimated change in outcomes for the same person in the same school associated with having one more AP. The results from these analyses suggest that having one additional AP is not associated with any significant differences in teachers' outcomes. For student outcomes, there are no measurable differences in a students' attendance rate, chronic absentee status, or likelihood of in-school suspension, but there is a positive relationship between the likelihood of receiving any out-of-

school suspensions and having one additional AP. Having one additional AP is associated with a .12 percent increase in the likelihood of having any out-of-school suspensions.

Table I.11 presents the results from models that estimate the association between APs' evaluation ratings to teacher and student outcomes, which addresses the third research question of this study. The models in Table I.11 restrict the sample to schools with one AP, so no changes to the number of APs occurs. The estimates on APs' evaluation ratings can be understood as the difference in outcomes for the same person in the same school associated with a one-point increase in APs' evaluation ratings. A more intuitive way of thinking about the estimates in Table I.11 is to think of them as the change in outcomes associated with having an AP that is rated one point higher than the average AP. A standard deviation in APs' evaluation ratings is approximately .58 points on the evaluation scale, so a one-point change in APs' evaluation ratings is a relatively large change. The results from Table I.11 suggest that there are no measurable differences in student outcomes associated with higher AP evaluation ratings. A one-point increase in the evaluation rating of a teacher's AP is associated with about a one-point reduction in the likelihood of a teacher turning over. There is also about a .08 SD increase in teachers' perceptions of relationships associated with a one-point increase in the AP's evaluation ratings. This result means that the same teacher while in the same school has more positive perceptions of their school climate and is less likely to turnover in years when they have an AP that has a higher evaluation rating.

Tables I.12-I.14 report the results for the student sub-sample analyses along race. These tables address the fourth research question. In each of these tables, Panel A reports the results for white students, Panel B for Black students, and Panel C for Latinx students. The sub-sample analyses for other race students, FRPL eligible students, and students with IEPs are in Appendix

Tables I.A4-I.A6. The results for these other student sub-samples do not suggest that there are many relationships that are systematically different across these sub-samples. Table I.12 replicates the analyses in Table I.9 for white, Black, and Latinx students. The results for these analyses suggest that most of the significant relationships between having an AP and student outcomes are concentrated among Black students. Having an AP is associated with a lower likelihood of receiving any in-school suspensions, increased attendance rates, and lower likelihoods of being chronically absent for Black students. White students have a slightly lower rate of chronic absenteeism, but they do not have a lower likelihood of in-school suspension or higher attendance rates as is found in the full sample of students. In the full sample having one more assistant principal is associated with a higher likelihood of receiving out-of-school suspensions and Table I.13 suggests that this relationship is strongest for white and Latinx students. Black and white students have a higher likelihood of receiving any in-school suspensions associated with having an AP, but Latinx students have a lower likelihood of receiving any in-school suspensions. Table I.14 presents the results for the sub-sample analyses for APs' evaluation ratings. While APs' evaluation ratings were not significantly related to student outcomes in the full sample, a one-point increase in APs' evaluation ratings is associated with an approximately .06 percent increase in white student's attendance rates.

I.6 Sensitivity Analyses

Since there is strong descriptive evidence that assignment of APs to schools is non-random, the validity of the estimated coefficients rests on whether the estimation strategy sufficiently accounts for potential sources of bias. One potential threat to the validity of these estimates is the potential of pre-treatment trends to bias the estimates. A hypothetical example of this type of

threat would be if APs are assigned to schools that the district knows are struggling with school climate. Given the choice between two possible schools that could qualify for an AP, districts systematically choose to assign APs to the school that is experiencing a downward trend in school climate. If those pre-trends continue to drive results after the AP is assigned to the school, the worse outcomes in periods after the AP arrives in the school would be attributed to the AP rather than the pre-treatment trends.

This threat of pre-treatment trends is most likely to bias the estimates of having an AP because this change in exposure to APs is most likely to occur in response to prior conditions in the school. The other two types of AP exposure are less likely to suffer from this pre-treatment trend because within schools the year-to-year change in the types of exposure have more variation and do not have a discernable relationship with the outcome variables in years prior to a change. Since APs are a large financial expenditure relative to other types of employees in a school, schools deciding to add or eliminate an AP position are not likely to make the decision lightly. By comparison AP evaluation ratings are likely to change at least a little each year and schools with multiple APs are more likely to experience more frequent changes in the number of APs. On average schools with one AP experience about .32 points change in their evaluation ratings each year, and schools with multiple APs experience an average .46 change in the number of APs each year. Schools with zero or one AP only experience an average change of .17 APs each year.

To test the extent to which the estimates in this study may be biased by pre-treatment trends, I subset the analytic sample for the first research question on the schools that experience a change in having an AP and have at least three years of data prior to the change. I then include those schools in all years before and after a change as long as that school does not experience

another change in having an AP. I then create a time variable for relative time that is centered on the year prior to a change in having an AP. The descriptive unadjusted mean outcomes before and after a change in having an AP are presented in Figures I.2-I.3. Figure I.2 presents the teacher outcomes and Figure I.3 presents the student outcomes. The first observation that stands out from these figures is that the teacher year-to-year outcomes seem to be more variable than the student outcomes. The second observation I take from these figures is that the student outcomes seem to exhibit stronger evidence of a sharper contrast in trends before and after a change in having an AP. Although these figures are illustrative, they do not account for any differences in students and teachers in years before and after a change in having an AP.

To account for differences in students and teachers, I estimate the same models as described in Equation 1 but substitute the *Has an AP* indicator for a vector of dummies for each relative year excluding the year prior to a change, and estimate the models separately for schools that gain an AP and schools that lose an AP.[5] I present these results graphically in Figures I.4 and I.5. These estimates are much less precise because they rely on a small sub-sample of schools and estimate separate relationships for each relative year. Although these standard errors are much larger, the point estimates can provide some suggestive evidence of any pre-trends or divergences in outcomes in years prior to a change. Figure I.4 suggests that there may have been a downward trend in perceptions of relationships prior to losing an AP that may have carried into post-change years. Figure I.4 also suggests that teachers were more likely to turnover in years prior to a change in APs, but the pattern in outcomes seems to flatten out in years after a change.

[5] I also estimate these models using the relative time dummy variables and interactions with an indicator for schools that lose an AP. These models that include both schools that gain an AP and lose an AP do not produce substantively different results.

These results suggest that, for teachers, the pre-existing trends in outcomes are most likely to influence teacher perceptions of relationships in schools experiencing a loss in APs.

Figure I.5 seems to suggest that there are not many deviations in years prior to a change in APs for student outcomes. The one outcome that seems to exhibit pre-treatment trends is having an out-of-school suspension for schools that lose an AP. It seems that students' likelihoods of receiving out-of-school suspensions were increasing prior to losing an AP and continued to rise after losing an AP. This result suggests that the estimate of changes in the likelihood of out-of-school suspensions associated with changes in having an AP may be more likely to suffer from bias due to pre-trends. It is important to note that all these analyses suggest that there are no measurable differences in years prior to a change in having an AP and any suggestions of bias are based on insignificant point estimates. Taking a conservative interpretation, these analyses seem to suggest the results for teachers' perceptions of community and students' likelihood of out-of-school suspensions are suspect and may suffer from bias. It is reassuring that most of the outcomes do not even have suggestive evidence of the presence of pre-treatment trends.

Another potential concern is that the relationship between APs and outcomes associated with school climate is driven primarily by changes in specific grade levels. The descriptive differences between schools with different levels of exposure to APs as reported in Tables I.2 – I.6 suggest that there may be differential relationships for exposure to APs across grade levels. To assess the extent to which the estimates of exposure to APs differs across grade levels, I conduct sub-sample analyses at each grade level (elementary, middle, and high schools). The results from these analyses are reported in Tables I.15-I.17. Table I.15 presents the results for having an AP and the only significant results are for students in elementary schools. Students in

elementary schools have lower likelihoods of receiving in-school suspensions and of being chronically absent when they have an AP as compared to when they do not have an AP. Table I.16 presents the sub-sample analyses across grade levels for having one additional AP beyond one. The results in Table I.16 suggest that there are generally worse outcomes for students in elementary and middle schools when they have one additional AP. Elementary school students have lower attendance rates and middle school students are more likely to have in-school-suspensions when they have one additional AP. However, the evidence is more mixed for high school students who experience changes in the number of APs that they have. High school students who have one additional AP have higher likelihoods of receiving in-school and out-of-school suspensions, but they have higher attendance rates and lower likelihoods of chronic absenteeism.

The positive outcomes associated with having a more highly rated AP are primarily concentrated among elementary schools. As is reported in Table I.16, teachers in elementary schools have lower likelihoods of turnover and students in elementary schools have lower likelihoods of receiving in-school suspensions and have higher attendance rates when they have APs with higher evaluation ratings. Middle school teachers have higher likelihoods of turning over when they have APs with higher evaluation ratings, but students in middle schools have higher attendance rates when they have APs with higher evaluation ratings. There are no significant differences in outcomes for high school teachers or students associated with increases in AP evaluation ratings. Overall, these analyses do not seem to invalidate the results, but they do suggest that they may be most applicable to students and teachers in elementary schools. This seems to be less of an issue of internal validity, and more of an issue of generalizability.

The next sensitivity analysis I conduct is to test the inferences made in the primary analyses when standard errors are estimated to account for clustering at the school level rather than at the unit-specific fixed effect level. I conduct these analyses because some methodological papers suggest that the appropriate level at which to cluster standard errors is at the level of treatment. These results are not presented in a table because almost all the significant relationships are no longer significant due to the larger standard errors when using the higher level of clustering. There are two exceptions to this overall finding: the relationship between having a higher rated AP and teacher turnover is still significant at a .1 level and the relationship between having an AP and chronic absenteeism for the sub-sample of Black students is still significant at a .1 level. These results suggest that when using the most generous inferences based on the smaller standard error estimates, there are generally positive associations between student and teacher outcomes and exposure to APs. Using the more conservative standard error estimates still suggests that there are positive outcomes associated with exposure to APs.

I.7 Discussion

The results from these analyses provide evidence to support the notion that APs play an important role in improving school climate. The results also suggest that the relationship between exposure to APs and school outcomes is complex and may be differential across different school contexts. The measurable differences in outcomes of having an AP as compared to not having an AP are primarily found in student outcomes. Students in the same school have lower likelihoods of receiving in-school suspension, higher rates of attendance, and lower likelihoods of being chronically absent when they have an AP, and these improved student outcomes are concentrated among Black and elementary school students. Having one additional AP seems to have no

measurable relationship with teacher outcomes, and generally worse outcomes for students, although the evidence suggests this negative relationship is highly dependent on the demographic characteristics of students. Having an AP with higher evaluation ratings is related to improved outcomes for teachers, and the sub-sample analyses suggest that there are some improvements to student attendance rates for white, elementary, and middle school students. In the grade level sub-samples, the relationship between APs' evaluation ratings and teachers' perceptions of the school community are no longer significant, but there is a lower likelihood of turnover for elementary school teachers but a higher likelihood of turnover for middle school teachers.

To generalize the results, having an AP seems to have the most positive differences at the elementary school level and especially for Black students. Having more highly rated APs has more positive differences for teachers and especially at the elementary school level. Having one additional AP produces mixed results for student outcomes but has the most positive relationships with student attendance at the high school level. These results suggest that any positive benefits of APs may depend on the school context. Additionally, these results suggest that APs have a relatively small influence on individual students and teachers, but in aggregate they may have more substantive impacts.

I.8 Limitations

This study aims to estimate the relationship between exposure to APs and school outcomes, but the descriptive findings suggest that generalizing the results to a broader set of school contexts is challenging. Schools that experience changes in the number of APs are not similar on observable characteristics to schools that do not experience changes in the number of APs. The results of this study are most applicable to the narrow set of schools that are similar on observables to

schools in this study. Another issue with the generalizability is that the student analyses are less likely to leverage middle school students' experiences because students are only in a middle school for a maximum of three years. The shorter time period in middle school means that there are fewer years for students to experience changes in exposure to APs while in the same school. However, schools in other states that experience changes in exposure to APs are likely to be similar to the schools in this study sample. Several states allocate APs using a similar school enrollment-based formula to Tennessee's (Kelly & Chesser, 2019).

Second, there may be unobserved changes or trends in schools that may be related to both changes to the APs in a school and the outcomes observed in this study. For instance, a district may choose to assign an AP to a school specifically because it is struggling with student discipline and attendance. It may be that a new AP is unable to make changes to the struggling school and the school continues to follow a trajectory of worsening student outcomes in the years after arriving. If the previous trend in student outcomes persists, changes to the number or quality of APs will be equated with the worsening student outcomes even though the changes to the AP position were not the cause of the changes to the student outcomes. Although my analyses could still suffer from this type of bias, my formal analyses of pre-treatment periods in schools that experience a change in having an AP do not seem to suggest that this bias is a concern in this study.

The person-by-school fixed effects included in my estimation strategy have the advantage of controlling for all time invariant characteristics of people (students or teachers) while they are in the same school. The estimator I employ and the results of the pre-trend analyses lend confidence that differences in outcomes associated with exposure to APs are truly reflective of the changes in exposure to APs rather than some other factor. The disadvantage of this approach

is that any people who are not in the same school before and after a change to APs effectively drop out of the estimates. This means that the results are most applicable to students and teachers who remain in the same school for many years.

Despite these challenges to the validity of this study, I believe that its methods and findings are unbiased. Though in theory there may be changes to schools that occur concurrently with changes in exposure to APs or some anticipatory response on the part of districts, the results do not suggest that this is a concern. I control for the factors that are most likely to be related to changes to APs in schools, so the chances that some other time-variant factor influenced these results is unlikely. This analysis may not be generalizable to a large set of schools, but the results are relevant to exactly the type of schools that will experience similar changes in exposure to APs. These results provide some evidence of the circumstances that may facilitate a positive impact on school climate from a change in AP exposure.

I.9 Conclusion

Despite these challenges to the generalizability of this study, it presents fresh evidence regarding the relationship between school leadership and school outcomes. The sample for this study may be stylized, but the characteristics of these schools are similar to schools in other contexts that are likely to experience changes to their APs. This study demonstrates that measurable relationships between exposure to APs and outcomes associated with school climate can be detected. Moreover, this study finds that there are many improvements in teacher and student outcomes associated with increases in the quantity and quality of APs. Taking the most conservative interpretation of these results, I would still conclude that exposure to APs has a significant and positive impact on the attendance of Black students and the turnover of teachers.

These two results alone suggest that APs can have a positive impact on school climate. These results do not validate the conceptual model proposed in this paper, but they suggest that the type of model proposed in this paper is possible.

These results also suggest that more thought and care could be placed into thinking about how to deploy APs to maximize their effectiveness. My analysis in this chapter estimates the contribution of APs to school climate outcomes as observed in practice, but there may be ways to further enhance the effectiveness of APs through policy interventions targeting APs. Future work should further explore the mechanisms by which APs influence school outcomes. An especially important question for future research is to explore how and why exposure to APs seems to have more concentrated positive impacts for specific populations of teachers and students.

Chapter I Figures

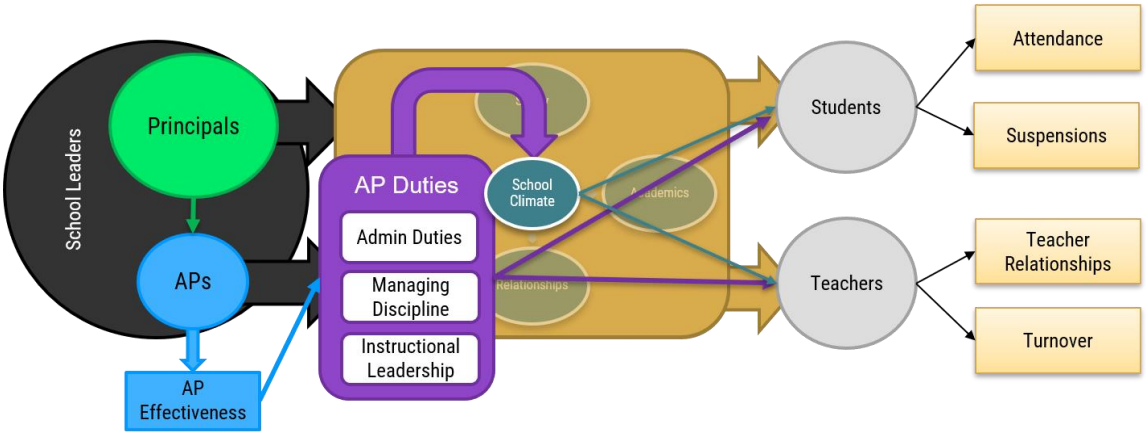
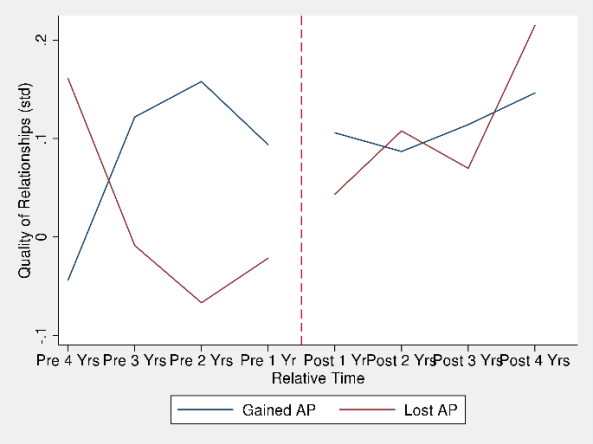
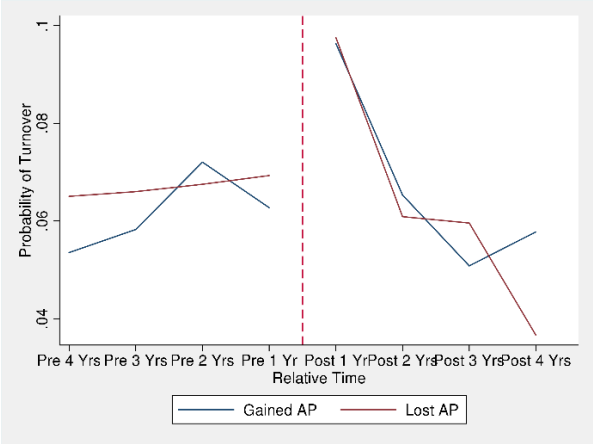


Figure I.1: Conceptual framework

Note: This conceptual framework shows the theoretical pathway between APs and student and teacher outcomes.



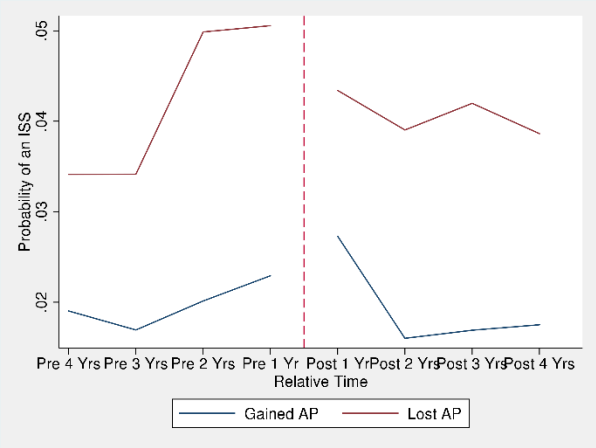
Teacher's Perceptions of Quality Relationships



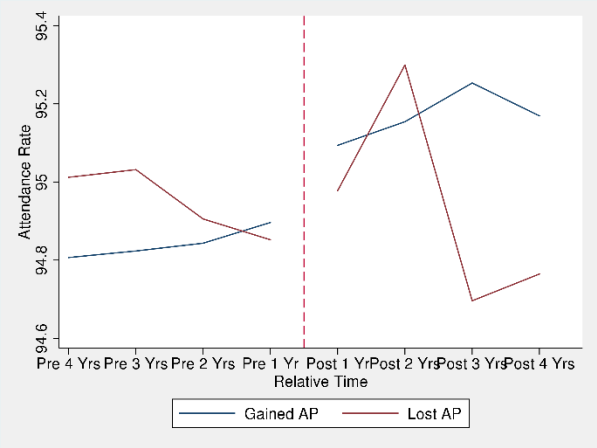
Teacher Turnover

Figure I.2: Average teacher outcomes before and after a change in having an AP

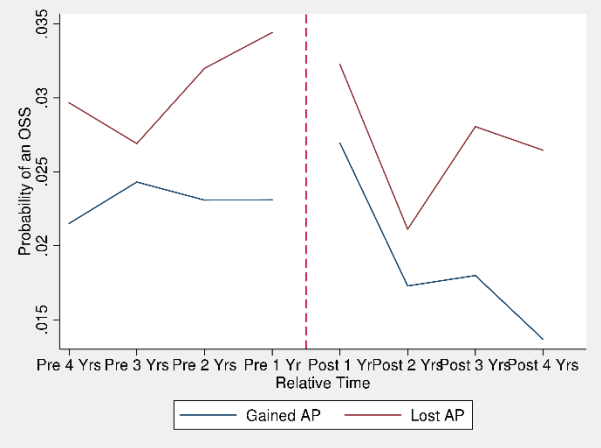
Notes: Gained AP refers to those schools that went from not having an AP to having an AP. There were 1472 teacher-by-year observation and 258 schools. Lost AP refers to those schools that went from having an AP to not having an AP. There were 1033 teacher-by-year observations and 180 schools.



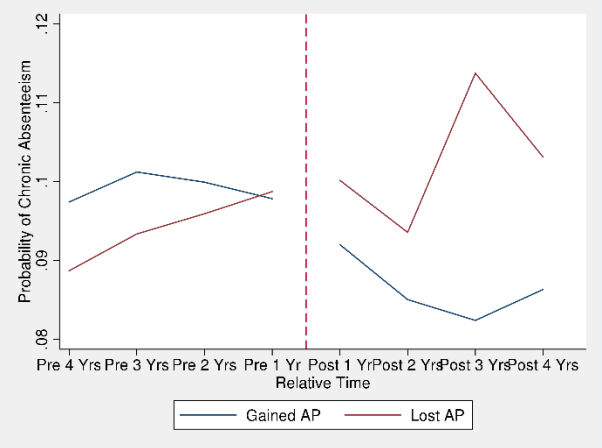
In-School-Suspension



Attendance Rate



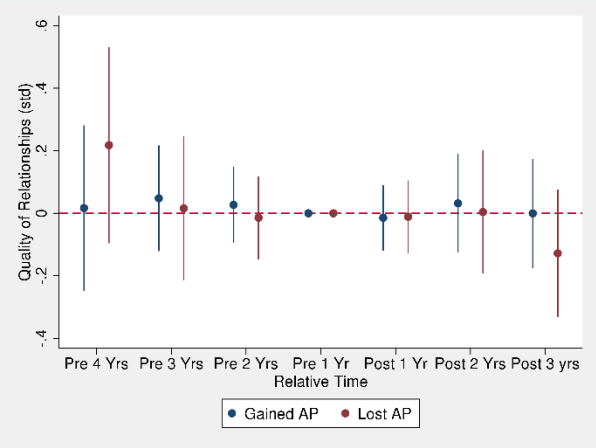
Out-of-School Suspension



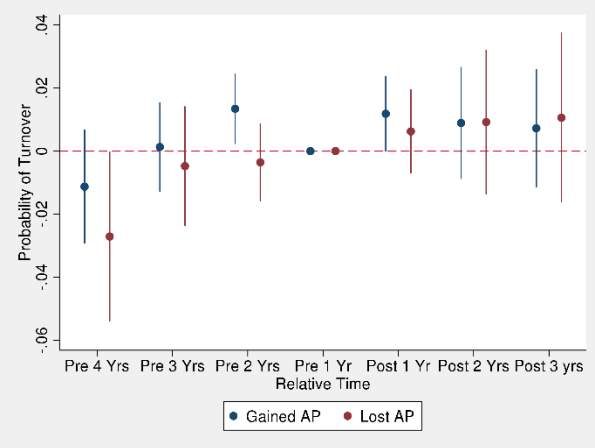
Chronic Absenteeism

Figure I.3: Average student outcomes before and after a change in having an AP

Notes: Gained AP refers to those schools that went from not having an AP to having an AP. There were 637,391 student-by-year observations and 258 schools. Lost AP refers to those schools that went from having an AP to not having an AP. There were 388,626 student-by-year observations and 180 schools.



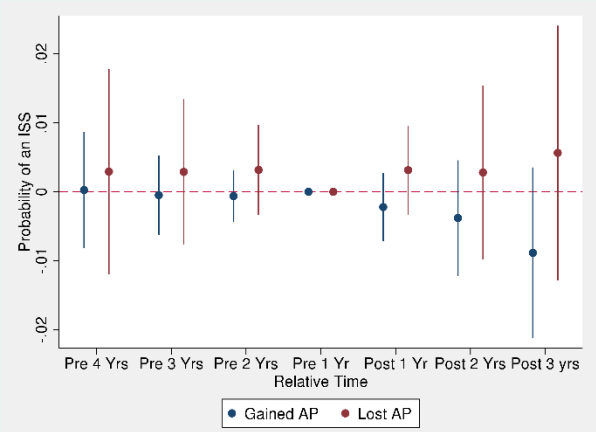
Teacher's Perceptions of Quality Relationships



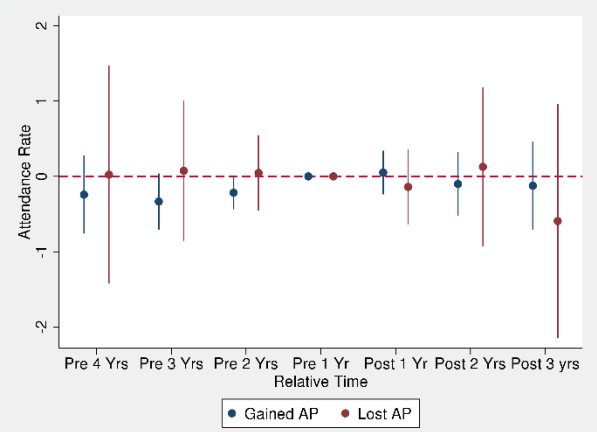
Teacher Turnover

Figure I.4: Coefficients for teacher outcomes regressed on relative time

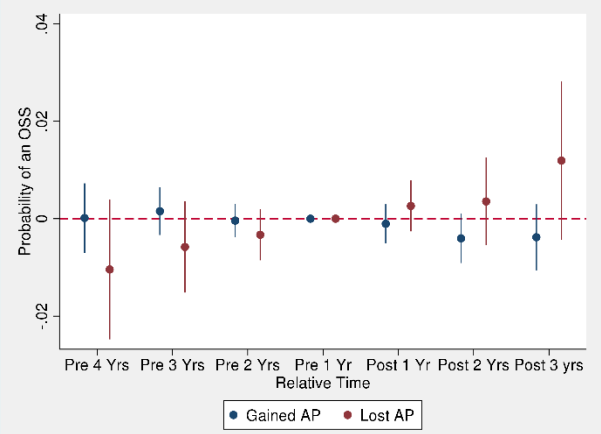
Notes: Gained AP refers to those schools that went from not having an AP to having an AP. There were 1472 teacher-by-year observation and 258 schools. Lost AP refers to those schools that went from having an AP to not having an AP. There were 1033 teacher-by-year observations and 180 schools.



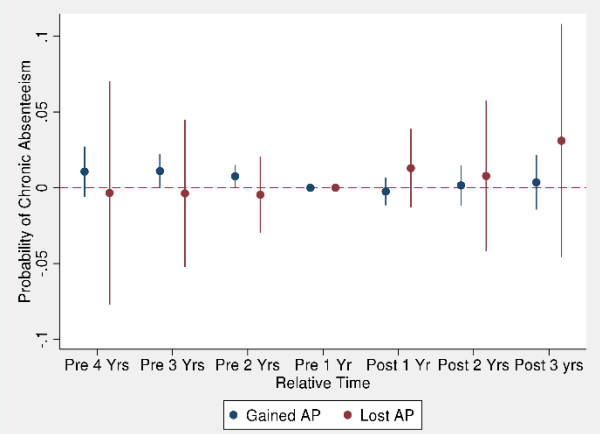
In-School-Suspension



Attendance Rate



Out-of-School Suspension



Chronic Absenteeism

Figure I.5: Coefficients for student outcomes regressed on relative time

Notes: Gained AP refers to those schools that went from not having an AP to having an AP. There were 637,391 student-by-year observations and 258 schools. Lost AP refers to those schools that went from having an AP to not having an AP. There were 388,626 student-by-year observations and 180 schools.

Chapter I Tables

Table I.1 Counts of schools with APs and schools that are funded for APs

	No AP	Has an AP	Total
Panel A: Any AP FTE Funding			
Not funded for an AP	3,569	3,080	6,649
	(40.58)	(35.02)	(75.6)
Funded for an AP	369	1,777	2,146
	(4.2)	(20.2)	(24.4)
Total	3,938	4,857	8,795
	(44.78)	(55.22)	(100)
Panel B: Funded for at least 1 AP FTE			
Not funded for an AP	3,710	4,357	8,067
	(42.18)	(49.54)	(91.72)
Funded for an AP	228	500	728
	(2.59)	(5.69)	(8.28)
Total	3,938	4,857	8,795
	(44.78)	(55.22)	(100)

Note: Percentages of cells reported in parentheses. Only the sample of schools with 0 or 1 APs was included.

Table I.2: Schools with an AP compared to schools without an AP

	No AP	Has AP	Sig.
Elementary	0.65	0.52	***
Middle	0.07	0.22	***
High	0.05	0.11	***
Urban	0.29	0.29	
Suburban	0.11	0.15	***
Town	0.14	0.17	***
Rural	0.45	0.38	***
Proportion Black	0.21	0.22	
Proportion Latinx	0.07	0.09	***
Proportion Other	0.03	0.03	***
Proportion FRPL	0.66	0.61	***
Proportion IEP	0.18	0.16	***
ADM	365	563	***
School by year observations	4080	5383	

Notes: The sample is restricted to schools with either zero or one AP.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.3: Schools that experience a change in AP exposure as compare to schools that do not

	Rest of Sample	Effective Sample	Sig.
Elementary	0.58	0.63	**
Middle	0.14	0.15	
High	0.08	0.08	
Urban	0.26	0.34	***
Suburban	0.14	0.11	*
Town	0.17	0.15	
Rural	0.43	0.41	
Proportion Black	0.19	0.25	***
Proportion Latinx	0.08	0.09	*
Proportion Other	0.03	0.03	
Proportion FRPL	0.62	0.65	***
Proportion IEP	0.17	0.16	
ADM	473	457	*
School by year observations	7646	1045	

Notes: The sample is restricted to schools with zero or one AP.
+ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table I.4: Schools that gain an AP compared to schools that lose an AP

	Lost an AP	Gained an AP	Sig.
Elementary	0.61	0.65	
Middle	0.15	0.15	
High	0.08	0.08	
Urban	0.34	0.34	
Suburban	0.09	0.12	
Town	0.13	0.16	
Rural	0.44	0.38	
Proportion Black	0.24	0.25	
Proportion Latinx	0.08	0.09	
Proportion Other	0.03	0.03	
Proportion FRPL	0.66	0.65	
Proportion IEP	0.16	0.16	
ADM	455	458	
School by year observations	446	599	

Note: The sample is restricted to schools that experience a change in the number of APs from the prior year and have zero or one AP.
+ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table I.5: Schools with one AP compared to schools with multiple APs

	One AP	Multiple APs	Sig.
Elementary	0.52	0.17	***
Middle	0.22	0.29	***
High	0.11	0.46	***
Urban	0.29	0.33	**
Suburban	0.15	0.23	***
Town	0.17	0.16	
Rural	0.38	0.28	***
Proportion Black	0.22	0.26	***
Proportion Latinx	0.09	0.09	
Proportion Other	0.03	0.04	***
Proportion FRPL	0.61	0.54	***
Proportion IEP	0.16	0.14	***
ADM	563	991	***
School by year observations	5383	3154	

Note: The sample is restricted to schools with one or more APs.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.6: Schools with low rated APs as compared to schools with high rated APs

	Low Eval AP	High Eval AP	Sig.
Elementary	0.50	0.49	
Middle	0.23	0.24	
High	0.12	0.12	
Urban	0.30	0.30	
Suburban	0.12	0.15	***
Town	0.17	0.15	
Rural	0.41	0.39	
Proportion Black	0.22	0.19	***
Proportion Latinx	0.10	0.07	***
Proportion Other	0.03	0.03	**
Proportion FRPL	0.63	0.53	***
Proportion IEP	0.16	0.14	***
ADM	560	586	***
School by year observations	2937	5670	

Note: APs with low evaluation ratings are defined as having a rating below a 4 and APs with high evaluation ratings are defined as having an evaluation rating of 4 or higher. The sample is restricted to schools with one AP.
+ $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table I.7: Distribution of outcomes for the three primary samples used in this study

	Mean	SD
Panel A: Schools with zero or one AP		
Teachers' Perceptions of Relationships	0.03	0.98
Teacher Turnover	0.14	0.35
Any ISS	0.04	0.20
Any OSS	0.04	0.19
Attendance Rate	94.91	6.63
Chronically Absent	0.10	0.30
Panel B: Schools with one AP		
Teachers' Perceptions of Relationships	0.00	0.98
Teacher Turnover	0.14	0.35
Any ISS	0.05	0.22
Any OSS	0.04	0.19
Attendance Rate	94.97	6.42
Chronically Absent	0.10	0.30
Panel C: Schools with one or more APs		
Teachers' Perceptions of Relationships	-0.04	0.98
Teacher Turnover	0.14	0.35
Any ISS	0.08	0.27
Any OSS	0.06	0.23
Attendance Rate	94.47	7.50
Chronically Absent	0.12	0.32

Note: Panel A is the sample of schools used to answer research question 1. Panel B is the sample of schools used to answer research question 2. Panel C is the sample of schools used to answer research question 2

Table I.8: Comparison of models estimating teachers' perceptions of relationships on having an AP

	(1)	(2)	(3)	(4)
Has an AP	-0.0545***	-0.0237+	-0.0243	-0.0192
	(0.0113)	(0.0136)	(0.0183)	(0.0193)
Covariates	X	X	X	X
School FE		X		
Person FE			X	
Person x School FE				X
Adj. R-Sq	0.05	0.19	0.52	0.56
N	65423	65423	65423	65423

Notes: Standard errors clustered at fixed effects level in columns 2-4. Standard errors in column 1 are Huber-White heteroskedastic robust estimates. Dependent variable is the factor scores for a teacher's perceptions of the quality of relationships in the school.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.9: Relationship between outcomes associated with school climate and having an AP

	Teacher Outcomes		Student Outcomes			
	Perceptions of Relationships	Turnover	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)	(5)	(6)
Has an AP	-0.0192	0.0031	-0.0013*	-0.0006	0.0286+	-0.0036***
	(0.0193)	(0.0031)	(0.0006)	(0.0006)	(0.0165)	(0.0010)
Adj. R-Sq	0.56	0.45	0.41	0.40	0.58	0.42
N	65420	167180	3576112	3576112	3576112	3576112

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. The sample is limited to schools with either zero or one APs.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.10: Relationship between outcomes associated with school climate and the number of APs

	Teacher Outcomes		Student Outcomes			
	Perceptions of Relationships	Turnover	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)	(5)	(6)
Number of APs	-0.0142	-0.0007	0.0009	0.0012**	0.0059	-0.0004
	(0.0126)	(0.0019)	(0.0005)	(0.0004)	(0.0130)	(0.0006)
Adj. R-Sq	0.57	0.44	0.39	0.40	0.56	0.43
N	76733	207989	4780327	4780327	4780327	4780327

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. The sample is limited to schools with one or more APs.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.11: Relationship between outcomes associated with school climate and evaluation ratings of APs

	Teacher Outcomes		Student Outcomes			
	Perceptions of Relationships	Turnover	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)	(5)	(6)
AP's Evaluation Rating	0.0841**	-0.0095**	-0.0013	0.0003	0.0272	-0.0006
	(0.0267)	(0.0037)	(0.0008)	(0.0007)	(0.0203)	(0.0011)
Adj. R-Sq	0.59	0.46	0.43	0.42	0.57	0.43
N	33013	86247	2200048	2200048	2200048	2200048

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. The sample is limited to schools with one AP.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.12. Sub-group analyses of student outcomes and having an AP by student race

	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)
Panel A: White				
Has an AP	-0.0007	-0.0006	0.0105	-0.0031**
	(0.0007)	(0.0006)	(0.0190)	(0.0011)
Adj. R-Sq	0.39	0.31	0.57	0.41
N	2545849	2545849	2545849	2545849
Panel B: Black				
Has an AP	-0.0045*	-0.0005	0.1374**	-0.0074**
	(0.0020)	(0.0023)	(0.0454)	(0.0026)
Adj. R-Sq	0.43	0.42	0.61	0.45
N	700049	700049	700049	700049
Panel C: Latinx				
Has an AP	-0.0009	-0.0006	0.0035	-0.0024
	(0.0016)	(0.0014)	(0.0464)	(0.0027)
Adj. R-Sq	0.37	0.32	0.58	0.40
N	337955	337955	337955	337955

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. Each panel presents the results for a sub-sample of students that identify as the race listed.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.13. Sub-group analyses of student outcomes and number of APs by student race

	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)
Panel A: White				
Number of APs	0.0011+	0.0012**	0.0166	0.0001
	(0.0006)	(0.0004)	(0.0155)	(0.0007)
Adj. R-Sq	0.36	0.31	0.56	0.42
N	3279847	3279847	3279847	3279847
Panel B: Black				
Number of APs	0.0026+	0.0006	0.0075	-0.0010
	(0.0015)	(0.0014)	(0.0296)	(0.0014)
Adj. R-Sq	0.41	0.40	0.59	0.47
N	996951	996951	996951	996951
Panel C: Latinx				
Number of APs	-0.0030+	0.0034*	-0.0609	-0.0012
	(0.0017)	(0.0014)	(0.0476)	(0.0019)
Adj. R-Sq	0.35	0.29	0.52	0.39
N	462867	462867	462867	462867

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. Each panel presents the results for a sub-sample of students that identify as the race listed.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.14. Sub-group analyses of student outcomes and AP evaluation ratings by student race

	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)
Panel A: White				
AP Evaluation Rating	-0.0003	0.0004	0.0570*	-0.0013
	(0.0009)	(0.0007)	(0.0241)	(0.0013)
Adj. R-Sq	0.41	0.32	0.56	0.42
N	1559921	1559921	1559921	1559921
Panel B: Black				
AP Evaluation Rating	-0.0039	0.0009	-0.0117	-0.0006
	(0.0024)	(0.0028)	(0.0516)	(0.0029)
Adj. R-Sq	0.46	0.44	0.59	0.47
N	415530	415530	415530	415530
Panel C: Latinx				
AP Evaluation Rating	0.0018	-0.0013	-0.0417	0.0006
	(0.0019)	(0.0017)	(0.0568)	(0.0029)
Adj. R-Sq	0.38	0.34	0.54	0.41
N	185990	185990	185990	185990

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. Each panel presents the results for a sub-sample of students that identify as the race listed.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.15: Sensitivity analyses – subgroup analyses by grade level for having an AP

	Teacher Outcomes		Student Outcomes			
	Perceptions of Relationships	Turnover	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Elementary						
Has an AP	-0.0151	0.0017	-0.0015**	-0.0008	0.0269	-0.0033**
	(0.0230)	(0.0037)	(0.0006)	(0.0006)	(0.0180)	(0.0011)
Adj. R-Sq	0.56	0.44	0.27	0.35	0.54	0.40
N	40178	104013	2220867	2220867	2220484	2220867
Panel B: Middle						
Has an AP	-0.0907	0.0095	-0.0038	0.0018	0.0829	-0.0054
	(0.0636)	(0.0096)	(0.0036)	(0.0029)	(0.0562)	(0.0034)
Adj. R-Sq	0.57	0.48	0.39	0.42	0.61	0.46
N	9207	23773	569258	569258	569258	569258
Panel C: High						
Has an AP	-0.0097	0.0039	0.0055	-0.0014	0.1256	0.0003
	(0.0820)	(0.0137)	(0.0046)	(0.0034)	(0.1344)	(0.0061)
Adj. R-Sq	0.60	0.48	0.32	0.36	0.64	0.43
N	4641	11980	270307	270307	270307	270307

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. The sample is limited to schools with zero or one APs.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.16: Sensitivity analyses – subgroup analyses by grade level for number of APs

	Teacher Outcomes		Student Outcomes			
	Perceptions of Relationships	Turnover	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Elementary						
Number of APs	-0.0321	-0.0035	-0.0002	0.0003	-0.0690***	0.0015
	(0.0319)	(0.0045)	(0.0006)	(0.0006)	(0.0199)	(0.0010)
Adj. R-Sq	0.55	0.44	0.34	0.38	0.54	0.43
N	28371	75505	1656359	1656359	1656359	1656359
Panel B: Middle						
Number of APs	-0.0330	0.0058	0.0060***	-0.0008	-0.0174	-0.0010
	(0.0289)	(0.0052)	(0.0015)	(0.0013)	(0.0251)	(0.0014)
Adj. R-Sq	0.56	0.44	0.41	0.43	0.62	0.47
N	16146	43171	1085879	1085879	1085879	1085879
Panel C: High						
Number of APs	-0.0215	-0.0029	0.0016+	0.0022***	0.0633**	-0.0019*
	(0.0166)	(0.0024)	(0.0008)	(0.0007)	(0.0201)	(0.0009)
Adj. R-Sq	0.60	0.45	0.37	0.38	0.58	0.44
N	23178	66216	1515233	1515233	1515233	1515233

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. The sample is limited to schools with one or more APs.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.17: Sensitivity analyses – subgroup analyses by grade level for AP evaluation rating

	Teacher Outcomes		Student Outcomes			
	Perceptions of Relationships	Turnover	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Elementary						
AP Evaluation Rating	0.0571	-0.0181***	-0.0010+	0.0007	0.0949***	-0.0018
	(0.0367)	(0.0049)	(0.0006)	(0.0008)	(0.0234)	(0.0013)
Adj. R-Sq	0.59	0.44	0.32	0.36	0.54	0.42
N	18980	50320	1069202	1069202	1069002	1069202
Panel B: Middle						
AP Evaluation Rating	0.0490	0.0221*	-0.0028	0.0016	0.1049+	-0.0019
	(0.0604)	(0.0108)	(0.0040)	(0.0031)	(0.0608)	(0.0034)
Adj. R-Sq	0.63	0.49	0.39	0.41	0.62	0.46
N	6030	15676	368749	368749	368749	368749
Panel C: High						
AP Evaluation Rating	0.0776	-0.0162	-0.0046	-0.0020	-0.0483	0.0067
	(0.0826)	(0.0123)	(0.0047)	(0.0036)	(0.1272)	(0.0059)
Adj. R-Sq	0.65	0.48	0.35	0.37	0.58	0.43
N	2955	7751	161492	161492	161492	161492

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. The sample is limited to schools with one AP.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Chapter I Appendix Figures

Elementary Assistant Principals are allocated according to the following schedule.

School Enrollment	Positions Allocated
Below 660	0.0
660 – 879	0.5
880 – 1,099	1.0
1,100 – 1,319	1.5
Above 1,319	2.0

EXAMPLE:

Enrollment	Positions
567	0.0
666	0.5
990	1.0
1,256	1.5
1,430	2.0

Secondary Assistant Principals are allocated according to the following schedule.

School Enrollment 9 – 12	Positions Allocated
Below 300	0.0
300 – 649	0.5
650 – 999	1.0
1,000 – 1,249	1.5
Above 1,249	2.0 (plus 1 for each additional 250 pupils rounded to nearest .5)

EXAMPLE:

Enrollment	Positions
280	0.0
555	0.5
875	1.0
1,200	1.5
1,589	3.0

The number of positions is multiplied by the state instructional salary unit cost as set by the annual appropriations bill to determine the total component support. For FY14 the state instructional salary unit cost is $40,447.

Figure I.A1: BEP funding formula for APs in Tennessee

KCS
KNOX COUNTY SCHOOLS

Teacher Funding - How Does it Work?

The Basic Education Program (BEP) is the funding formula through which state education dollars are generated and distributed to Tennessee schools. The BEP funds are what the state defines as sufficient to provide a basic level of education for students.

The district staffs our schools at a more generous level in many cases than the BEP funds provide. For example, at the high school level, the BEP funds a student teacher ratio of 30:1. In contrast, the district funds 13 out of our 16 high schools at a ratio of 25:1 or lower.

Basic Education Program (BEP) Staffing Ratios

These are the state average class size requirements

Elementary		Middle		High	
K - 3	1:20	**6**	1:25	**9 - 12**	1:30
4 - 5	1:25	**7 - 8**	1:30		

**** Knox County Schools receives $27,984 per position (based on the county's fiscal capacity index)***

2019-20 Knox County Schools Staffing Ratios

These are the ratios used to allocate teaching positions in the district

Elementary		Middle		High	
K - 3	1:20	*At-Risk < 50%*		*At-Risk < 15%*	
4 - 5	1:25	**6**	1:25	**9 - 12**	1:27
** At-Risk adjustments made by GLD and Staff*		**7 - 8**	1:30	At-Risk > 15% < 45%	
		At-Risk > 50%		**9 - 12**	1:25
		6	1:20	At-Risk > 45%	
		7 - 8	1:25	**9 - 12**	1:20

Figure I.A2: Knoxville County Schools' adjustments to state allocation formulas

FY19 SBB Formula			
Weights	**ES**	**MS**	**HS**
Base Weight (1.0)*	$4,600		
Grade Weight	.10	.05	
Prior Academic Performance (Poverty as a proxy in ES)	.10	.10	.05
English Learners	.24		
Poverty	.05		
Special Education	Varies by Option Type		
Adjustments	Small Schools/Hold-Harmless		

Figure I.A3: Metro Nashville Public Schools' budget allocation formula

Student Need	Rationale
Base Weight	• Base Weight-- $3,530 for all K12 General Education Setting Students • Students with Disabilities (SWD) Increment-- $825 or 0.24 for all SWD self-contained students -- These students do not receive the full base weight because their instructional resources are locked. This means that schools will continue to receive Special Ed teachers/TAs from the Special Ed department, and these students only need an "incremental" amount to cover their share of the school's administrative, operational, and other schoolwide services.
Grade Weight: **K-5 with emphasis on K-2**	• Grade Level-- $1,059 or 0.3 for K-2, $706 or 0.2 for 3-5 • Based on the resources that have been unlocked to schools, ES need a slightly higher weight to cover their lower class size requirements. Additionally, many of the locked resources (i.e., athletics, CTE, security, custodial, etc.) are places where SS are likely to get a higher share than ES. • K-2 is weighted more because literacy is critical district focus area and highly predictive of future outcomes. • DATA USED: 19-20 Projected enrollments by grade
Incoming **Student Performance (High and Low)**	• *Incoming* student performance (high and low) – $353 or 0.1 • Student performance is one of the most important indicators of student need at a school. • Note that this will be calculated as incoming student performance (i.e., in MS/HS, we will look at the TNReady performance of its incoming 6th graders and 9th graders when they were 5th graders and 8th graders respectively, and use those %s to determine need) DATA USED: ES/K-8 – Because incoming performance data is not available, • Low Performance: 17-18 TNReady Below/Approaching Proficient % of 4th graders in 18-19 enrollment year • High Performance: 17-18 TNReady Advanced % of 4th graders in 18-19 enrollment year MS/HS/6-12 • Low Performance: 17-18 TNReady Below/Approaching Proficient % of Incoming 6th/9th graders in 18-19 enrollment year • High Performance: 17-18 TNReady Advanced % of Incoming 6th/9th graders in 18-19 enrollment year
Mobility	• Mobility – $1,412 or 0.4 • Schools with highly mobile populations have greater levels of need. We will weight mobility instead of poverty because mobility is highly correlated to poverty but provides a more nuanced look at need. • DATA USED: Mobility rate (defined by number of students who transferred into the school after 20th day divided by number of students who ever attended the school, excluding students who attended for less than 1 week) calculated using year-end SY17-18 data.
Poverty	• Direct Certified - $353 or .10 • ELL - $88 or .03 (please note that you will continue to receive ELL resources (e.g. ESL Teachers) outside of SBB. This small weight is meant to alleviate imperfections in the measurement of poverty for the poverty weight). • DATA USED: The number of students eligible for the Poverty weight is calculated using the 2018-19 Direct Certified %

Figure I.A4: Shelby County Schools' budget allocation formula

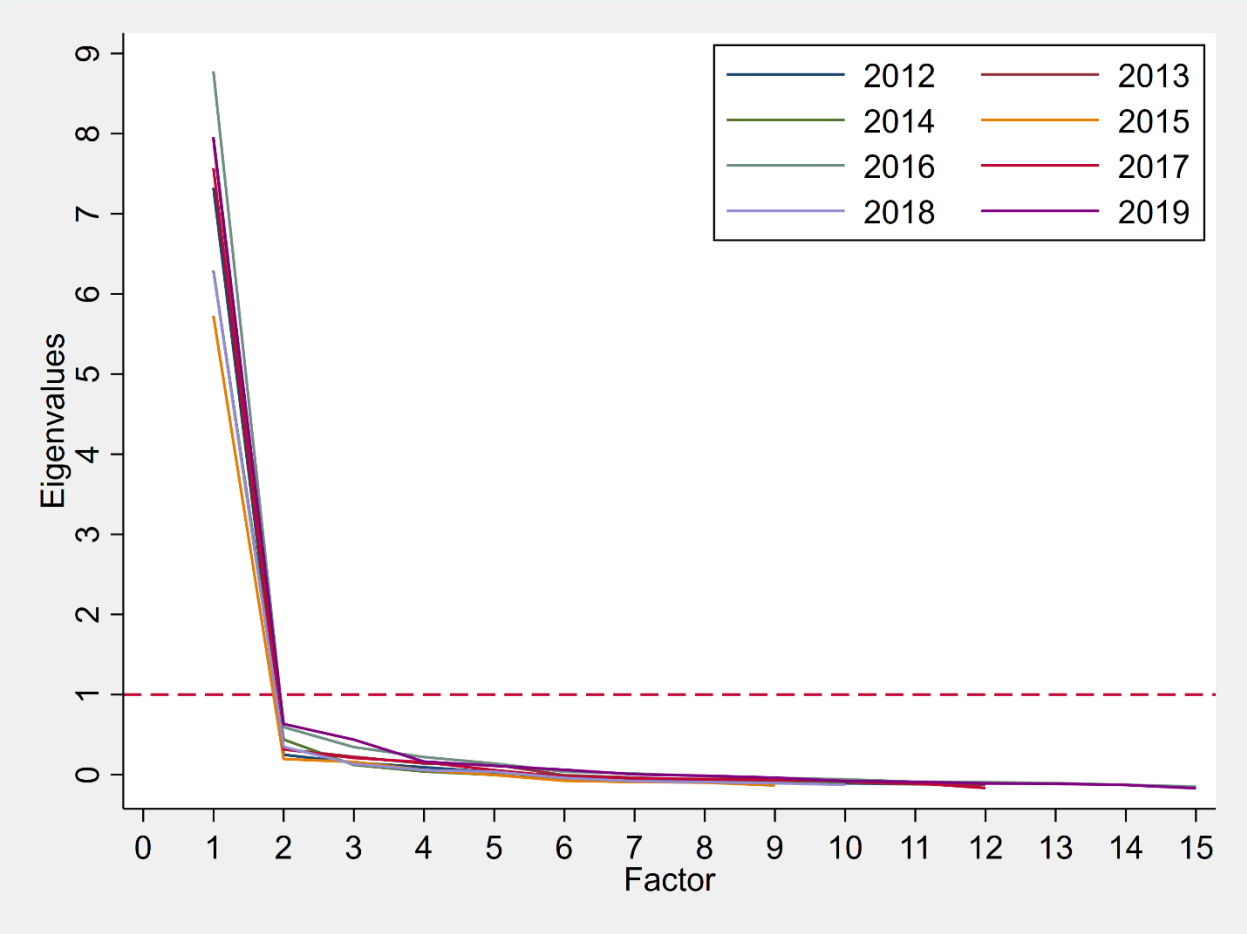


Figure I.A5 Scree plot of eigenvalues for factor analysis conducted for each year

Notes: This figure plots eigenvalues on the number of factors. A horizontal line is drawn to show how many factors are above and below the rule of thumb threshold of an eigenvalue of 1.

Table I.A1: Survey questions and years included

Question	2012	2013	2014	2015	2016	2017	2018	2019
Adults treat students with respect						Y	Y	
I feel appreciated for the job that I am doing			Y	Y		Y	Y	Y
I feel supported by other teachers at this school								Y
I like the way things are run in this school	Y	Y	Y			Y	Y	Y
I receive the supports needed to teach students of all cultures					Y			
I would recommend this school to parents seeking a place for their child								Y
Leaders are adequately visible and available to address needs					Y	Y	Y	Y
Leaders consistently support the school staff				Y				
Leaders facilitate teachers working together	Y	Y	Y		Y			
Leaders make an effort to address staff concerns				Y	Y	Y		
Leaders praise teachers that perform well	Y	Y	Y					
Leaders seek to understand staff needs					Y	Y	Y	Y
Leaders support risk-taking	Y	Y	Y					
Leaders trust the judgment of teachers	Y	Y	Y					
Leaders value teachers' ideas	Y	Y	Y					
Most of my colleagues share my beliefs				Y	Y			
Our school staff is a learning community								Y
Staff at this school have an effective process for making group decisions							Y	
Staff at this school have an effective process for solving problems								Y
Students in my school are often threatened and bullied								
Students treat adults with respect						Y		Y
Teacher are informed of current issues in school	Y	Y	Y					
Teacher involvement in decision making is serious	Y	Y	Y					
Teachers are encouraged to participate in school leadership roles				Y	Y	Y	Y	Y
Teachers are encouraged to share ideas	Y	Y	Y					
Teachers are involved in decision making	Y	Y	Y					
Teachers are rewarded for experimenting	Y	Y	Y					
Teachers have an appropriate level of influence on decision making				Y	Y	Y		
Teachers have opportunities for shared leadership					Y			
Teachers serve a major role in setting priorities					Y			
The staff at this school feels generally satisfied	Y	Y	Y	Y		Y	Y	Y
The staff feels comfortable raising issues				Y	Y	Y	Y	Y
There is an atmosphere of trust				Y	Y	Y	Y	Y
This school fosters appreciation of all cultures					Y			
Teacher response rate[1]	25	37	42	57	48	56	58	62

Note: [1]Teacher response rates as reported by the Tennessee Education Research Alliance https://www.tn.gov/content/dam/tn/education/data/2019-survey/Survey_Report.pdf

Table I.A2: Instrumental variables analysis – Instrumenting AP FTEs under BEP for number of APs

	Teacher Outcomes		Student Outcomes			
	Perceptions of Relationships	Turnover	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Schools with 0 or 1 AP						
Has an AP	-0.377	0.046	0.1320*	0.1015*	-1.5761	0.0262
	(0.895)	(0.037)	(0.0572)	(0.0458)	(1.4873)	(0.0426)
First Stage F	1.69	5.25	6.40	6.40	6.41	6.40
Chi Sq.	444.23	3376.24	2311.07	1983.07	7032.92	8201.49
df M	29	29	39	39	39	39
N	66539	171369	3523600	3523600	3523600	3523600
Panel B: All Schools						
Number of APs	-0.846*	0.005	0.3032	0.1271	-7.4501	0.2522
	(0.332)	(0.013)	(0.6936)	(0.2737)	(12.8303)	(0.4527)
First Stage F	7.87	7.86	0.21	0.21	0.21	0.21
Chi Sq.	268.01	6034.67	1456.40	2242.90	4801.84	6432.93
df M	29	29	39	39	39	39
N	107235	284874	6094574	6094574	6094574	6094574

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates and year fixed effects. The sample in Panel A is limited to schools with one AP, and the sample in Panel B includes all schools.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.A3: Reliability measures of teachers' perceptions of relationships

	2012	2013	2014	2015	2016	2017	2018	2019	Average
Inter-Item Correlations	0.46	0.42	0.43	0.39	0.34	0.36	0.37	0.24	0.38
Cronbach's Alphas	0.95	0.96	0.94	0.94	0.95	0.95	0.94	0.93	0.94
Number of Items Included	12	12	13	9	13	12	10	13	11.75

Table I.A4: Sub-group analyses of student outcomes and having an AP by student race

	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)
Panel A: Other Race				
Has an AP	0.0026	-0.0028	0.1449	-0.0068
	(0.0034)	(0.0030)	(0.1065)	(0.0053)
Adj. R-Sq	0.39	0.40	0.56	0.43
N	4780327	4780327	4780327	4780327
Panel B: Non-FRPL				
Has an AP	-0.0000	0.0005	0.0306	-0.0024*
	(0.0008)	(0.0006)	(0.0234)	(0.0012)
Adj. R-Sq	0.41	0.40	0.65	0.44
N	1408578	1408578	1408377	1408377
Panel C: FRPL				
Has an AP	-0.0021*	-0.0008	0.0148	-0.0039**
	(0.0010)	(0.0009)	(0.0244)	(0.0015)
Adj. R-Sq	0.41	0.39	0.58	0.41
N	2167534	2167534	2167534	2167534
Panel D: No IEP				
Has an AP	-0.0010	-0.0006	0.0279	-0.0035***
	(0.0007)	(0.0006)	(0.0177)	(0.0010)
Adj. R-Sq	0.40	0.40	0.59	0.42
N	2998562	2998562	2998562	2998562
Panel E: Has IEP				
Has an AP	-0.0019	-0.0006	0.0550	-0.0039
	(0.0020)	(0.0020)	(0.0533)	(0.0031)
Adj. R-Sq	0.42	0.41	0.55	0.41
N	577550	577550	577550	577550

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. Each panel presents the results for a sub-sample of students that identify as the race listed.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.A5: Sub-group analyses of student outcomes and AP evaluation ratings by student race

	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)
Panel A: Other Race				
AP's Evaluation Rating	-0.0001	-0.0003	0.0292	0.0043
	(0.0037)	(0.0032)	(0.1090)	(0.0052)
Adj. R-Sq	0.38	0.34	0.54	0.40
N	223710	223710	223710	223710
Panel B: Non-FRPL				
AP's Evaluation Rating	-0.0001	-0.0007	0.0681*	-0.0014
	(0.0010)	(0.0008)	(0.0299)	(0.0014)
Adj. R-Sq	0.42	0.39	0.63	0.44
N	771253	771253	771163	771163
Panel C: FRPL				
AP's Evaluation Rating	-0.0022+	0.0012	-0.0218	0.0009
	(0.0013)	(0.0012)	(0.0305)	(0.0018)
Adj. R-Sq	0.42	0.39	0.56	0.41
N	1059408	1059408	1059408	1059408
Panel D: No IEP				
AP's Evaluation Rating	-0.0017*	-0.0000	0.0259	-0.0002
	(0.0008)	(0.0007)	(0.0213)	(0.0012)
Adj. R-Sq	0.42	0.39	0.57	0.43
N	1544217	1544217	1544217	1544217
Panel E: Has IEP				
AP's Evaluation Rating	0.0013	0.0026	0.0794	-0.0038
	(0.0026)	(0.0025)	(0.0705)	(0.0038)
Adj. R-Sq	0.45	0.42	0.53	0.42
N	286444	286444	286444	286444

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. Each panel presents the results for a sub-sample of students that identify as the race listed.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table I.A6: Sub-group analyses of student outcomes and number of APs by student race

	ISS	OSS	Attendance Rate	Chronic Absenteeism
	(1)	(2)	(3)	(4)
Panel A: Other Race				
Number of APs	0.0010	-0.0002	0.0338	-0.0045*
	(0.0020)	(0.0016)	(0.0540)	(0.0022)
Adj. R-Sq	0.35	0.30	0.52	0.39
N	477919	477919	477919	477919
Panel B: Non-FRPL				
Number of APs	0.0003	0.0005	-0.0139	0.0004
	(0.0006)	(0.0005)	(0.0153)	(0.0007)
Adj. R-Sq	0.39	0.40	0.64	0.42
N	2096375	2096375	2095913	2095913
Panel C: FRPL				
Number of APs	-0.0003	0.0013	0.0229	-0.0011
	(0.0010)	(0.0008)	(0.0225)	(0.0011)
Adj. R-Sq	0.39	0.39	0.56	0.42
N	2552334	2552334	2552334	2552334
Panel D: No IEP				
Number of APs	0.0008	0.0009+	-0.0077	-0.0005
	(0.0006)	(0.0005)	(0.0135)	(0.0006)
Adj. R-Sq	0.39	0.38	0.57	0.43
N	3976506	3976506	3976506	3976506
Panel E: Has IEP				
Number of APs	0.0013	0.0035*	0.0915*	-0.0001
	(0.0018)	(0.0016)	(0.0459)	(0.0020)
Adj. R-Sq	0.42	0.42	0.54	0.43
N	672203	672203	672203	672203

Notes: Standard errors clustered by person by school in parentheses. Dependent variables are listed above model numbers. All models include covariates, year fixed effects, and person by school fixed effects. Each panel presents the results for a sub-sample of students that identify as the race listed.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

CHAPTER II

Determinants of Assistant Principals' Work:

School organizational factors and their relationship to assistant principals' time use

There is a well-established research base on the way education personnel allocate their time and the factors related to how educators organize their work (E. B. Goldring et al., 2019; Grissom et al., 2013; Metzker, 2003; Stallings, 1980; Vannest & Hagan-Burke, 2010). This line of research on how educators allocate their time has largely focused on teachers and principals, with little attention to APs (Oleszewski et al., 2012; Sun, 2012). Although APs' impact on schools is likely to be more indirect as compared to teachers and even principals, prior research suggests that APs have the potential to influence several important school outcomes (Carpenter et al., 2017b; Keesor, 2005; Madhlangobe & Gordon, 2012). Moreover, the first chapter in this dissertation suggests that teachers' and students' outcomes are related to the quantity and quality of their APs. Understanding APs' work may reveal the mechanisms by which APs might impact schools.

Research on APs is generally thin and quantitative research on APs represents a small proportion of the available studies. Of the few studies that use quantitative methods to study APs, the majority explore the duties and roles of APs in schools. However, the available research on the work of APs is largely limited to describing what APs do in their jobs, and the most common method of measuring APs' duties has been to ask them to rank order tasks by the amount of time devoted to those tasks (Glanz, 1994; Kwan & Walker, 2008; Sun, 2012). Few of these studies extend beyond describing the work of APs to think about how their work may be shaped by personal and contextual factors, especially how the organizational characteristics of schools may shape APs' work. Research on teachers' and principals' time allocation suggests

that context and personal characteristics matter for how they organize their work (E. B. Goldring et al., 2008; Grissom et al., 2013; Horng et al., 2010; May et al., 2012; Spear-Swerling & Zibulsky, 2014; Stallings, 1980; Vannest & Hagan-Burke, 2010). The available research on APs has found some suggestive evidence that personal characteristics such as gender, race, and age are related to differences in how APs allocate their time (Hausman et al., 2002; McClellan & Casey, 2015; Walker & Kwan, 2009). Moreover, prior research has suggested that grade level and school size shape how APs allocate their time (Oliver, 2005; Walker & Kwan, 2009).

Most of these studies of APs' work have used a single year of surveys or interviews as their data source, limiting the generalizability of their findings (Chen et al., 2000; Glanz, 1994; Hausman et al., 2002; Koru, 1993; Lochmiller & Karnopp, 2016; McClellan & Casey, 2015; Oliver, 2003). These studies also lack a consistent method for measuring how assistant principals allocate their time to tasks, with most studies using a ranking of tasks (e.g., Glanz, 1994; Sun, 2012) and others asking APs to indicate the frequency with which they engage in a task (Hausman et al., 2002). None of these studies examine the relative strength of the relationship between contextual factors and time allocation of APs. Walker and Kwan (2009) include multiple contextual factors within the same regression models that have APs' time allocation as the outcome, but their study was conducted in Hong Kong, which may not generalize to the US context.

Education leadership research needs more rigorous studies of APs' time allocation to push beyond straightforward descriptions. The field does not have any research that explores why APs might allocate their time in the ways described in prior research. Are there patterns in how APs allocate their time? What contextual factors are related to different patterns in time allocation? This study intends to address this gap in the literature by estimating the relationship

between multiple factors and how APs allocate their time. It addresses the following research questions: How do APs allocate their time across different school contexts in Tennessee? What is the relationship between APs' personal and school characteristics and how APs allocate their time? What is the relationship between having support staff in a school and how APs allocate their time? What is the relationship between principals' characteristics and how APs allocate their time? What is the relationship between how other school leaders allocate their time and how APs allocate their time?

This study is especially important when considering AP is the most numerous school leadership position in the US[6], and most principals work as APs before entering the principalship (Farley-Ripple et al., 2012; Folsom et al., 2015; Fuller et al., 2016; Hollingworth & Dude, 2009). If the duties that APs perform as school leaders are important to the work of schools (Clayton & Goodwin, 2015; Marshall & Hooley, 2006; Williams et al., 2020), and APs' work experiences help to prepare them for the principalship (Bastian & Henry, 2015; Clark et al., 2009; Portin et al., 2003), then research should explore under what conditions APs focus on different duties. Furthermore, research suggests that it may be more accurate to define leadership in organizations as a series of roles and functions rather than as people in specific positions (Firestone, 1996; Freeston, 1987; Hausman & Goldring, 2001; Pitner, 1986, 1988).

In the following sections of this chapter, I review the relevant research on educators' time allocation and provide a description of APs in Tennessee. Then I provide an overview of the data from Tennessee employed in this study and the methods used to analyze those data. Then this chapter turns to a presentation of the results from the analyses of the factors that are associated with how APs allocate their time and a discussion of the implications of these findings. I

[6] According to the 205-2016 NTPS survey, there were 183,671 APs and principals across the US, and 90,410 of those staff were principals. See https://nces.ed.gov/programs/digest/current_tables.asp

conclude this chapter with a section on the limitations of this study, some concluding thoughts, and some directions for future research.

II.1 Literature Review

II.1.1 How school personnel allocate their time

Empirical studies of educators' time allocation have had a consistent presence in the education research landscape since the 1970s when researchers started to observe and analyze data related to how teachers structure their instructional time (Stallings, 1980). The research on how teachers organize their instructional time reveals that much of it is spent on the core duty of teaching students, but it also reveals that there is considerable variation in the specific tasks that teachers would engage in during their instruction (Berliner, 1990; Metzker, 2003; Stallings, 1980). Researchers studying the area of teacher time allocation finds that teachers who focus more of the classroom time on actively engaging students in activities related to valued educational outcomes tend to have higher levels of student achievement (Berliner, 1990; Fisher et al., 1981). Although the magnitude of the relationships between time use on academic learning and student achievement is relatively small (Fisher et al., 1981), the significant findings suggest that there are patterns in the instructional behaviors of effective teachers that are related to improvements in student test scores. In more recent years, fewer studies have examined how teachers allocate their time in favor of studying differences in teacher quality more broadly defined (Bold et al., 2017). This shift towards teacher quality from teacher time allocation may in part be due to critiques that using time as a measure of instruction does not account for the quality of the curriculum or the instructor (Berliner, 1990).

A similar line of research on principals' time allocation has grown in the last twenty years as the field finds mounting evidence that principals are important to student achievement (Leithwood & Jantzi, 2009; Liebowitz & Porter, 2019). Although an emphasis on large quantitative studies of principals' time allocation has grown in recent years, research on the duties of principals have their roots in qualitative studies that helped to expose the complexity of principals' work (Kmetz & Willower, 1982; Portin et al., 2009; Wolcott, 2003). Recent quantitative studies of how principals organize their work have confirmed two major findings of prior qualitative studies: principals' work is comprised of a diverse set of tasks and their work is shaped by the context of their schools (Camburn et al., 2010; E. B. Goldring et al., 2008; Grissom et al., 2013; Horng et al., 2010; Huang et al., 2018). Another important finding to arise from these recent quantitative studies of principals' time is that principals spend a large proportion of their time on managing the organization of schools despite a rhetorical emphasis in research on principals as instructional leaders of schools (Camburn et al., 2010; E. B. Goldring et al., 2008; Grissom et al., 2013; Huang et al., 2018). For example, Goldring et al. (2008) find that principals spend about eight hours a week on duties related to instructional leadership and almost ten hours a week on student affairs and five hours on personnel related tasks. Camburn et al. (2010) find that principals in an average day spend the most time on personnel issues and less time on leading the instructional activities in the school. Horng et al. (2010) find that principals spend the most time on managing students and organizational maintenance activities like managing the school's budget, and Grissom et al. (2013) find that principals only spend an average of 12.7 percent of their time on instructional activities. Huang et al. (2018) find that principals spend much of their time focusing on keeping an orderly environment and monitoring

progress and much less time on building the capacity of teachers through activities like mentoring and professional development.

II.1.2 What factors are related to how school personnel allocate their time

In addition to describing how school personnel spend their time, studies of time allocation in education have identified several personal and organizational characteristics that are related to differences in time allocation among school personnel (E. B. Goldring et al., 2008; Grissom et al., 2013; Stallings, 1980; Vannest & Hagan-Burke, 2010). For teachers, the number of years since they were certified and their knowledge of current research are related to how much time they spend on research based instructional practices (Spear-Swerling & Zibulsky, 2014). Vannest and Hagan-Burke (2010) find that how special education teachers allocate their time is related to the structure of their instructional duties (e.g., pull-out, resource, co-teaching). Several studies find that teachers' practices vary across different levels of school achievement (Stallings, 1980; Virgilio et al., 1991). School leadership and how they structure the school climate seem to play especially important roles in shaping the behaviors of teachers (Bryk et al., 2010; Julie Cohen & Brown, 2016; Hitt & Tucker, 2016; Smylie, 1988).

Prior research suggests that school leaders, specifically their leadership behaviors, are critical to quality instruction in schools (Coelli & Green, 2012; Hitt & Tucker, 2016; Leithwood et al., 2004; Liebowitz & Porter, 2019; Smylie, 1988). Although some leadership behaviors are related to improvements in student outcomes, research also finds that principals' time is often constrained by contextual factors. Several studies find that the school size is related to how school leaders allocate their time (E. B. Goldring et al., 2008, 2019; Grissom et al., 2013; Grissom, Loeb, et al., 2015). Leaders of smaller schools tend to spend more time on instructional

leadership tasks and less on internal relations (Grissom et al., 2013; Grissom, Loeb, et al., 2015). Grade level is another important contextual factor related to principals' time allocation in several studies (E. B. Goldring et al., 2008, 2019; Grissom et al., 2013; Grissom, Loeb, et al., 2015; Huang et al., 2018). These studies find that high school principals tend to spend more time on organizational management tasks, and elementary school principals spend more time on instructional tasks (E. B. Goldring et al., 2008; Grissom et al., 2013; Grissom, Loeb, et al., 2015).

The demographic characteristics of the student population also have a significant relationship with principals' time allocation across multiple studies. Grissom et al. (2013) finds that principals who spend more time on instructional leadership lead schools that have more Black and free and reduced price lunch eligible (FRPL) students, and Grissom et al. (2015) find that principals who lead schools with more FRPL students spend less time on tasks related to organizational management. Goldring et al. (2008) and Huang et al. (2018) find conflicting evidence about the diversity of tasks that principals spend their time on in schools with more disadvantaged student populations. Goldring et al. (2008) find that principals in schools with less disadvantaged student populations tend to spend their time on a wide range of activities while Huang et al. (2018) find the opposite pattern. Studies also find that principals' time allocation is related to student achievement. A few studies find that principals of schools with high achievement growth tend to spend less time on instructional activities and more time on tasks related to managing the organizational aspects of the schools (Grissom et al., 2013; Horng et al., 2010; May et al., 2012). However, Grissom et al. (2013) find that principals who spend more time specifically coaching and evaluating teachers tend to lead schools with higher achievement growth in math, and Goldring et al. (2008) find that schools with principals who spend their time

on instructional leadership tend to have higher levels of academic press. These studies seem to suggest that there is an important relationship between contextual factors and the duties that occupy a principal's time.

Another important contextual factor in determining how school leaders allocate their time is the local policy environment (Portin et al., 2003). One type of policy that research finds has had a significant role in shaping how school leaders spend their time in recent years is teacher evaluation policy (Cannata et al., 2017; E. Goldring et al., 2015; E. B. Goldring et al., 2019; Lochmiller & Mancinelli, 2019; Neumerski et al., 2018). These studies find that the implementation of multiple measure high stakes teacher evaluation policies in recent years has resulted in school leaders spending much more of their time on evaluating the instruction of teachers (Lochmiller & Mancinelli, 2019; Neumerski et al., 2018). As evaluation becomes an important part of the role of principals, prior studies find some evidence of principals shifting some of their organizational management duties to assistant principals (E. B. Goldring et al., 2019; Lochmiller & Mancinelli, 2019; Neumerski et al., 2018). Just as principals have adapted to changes to teacher evaluation, APs' duties and tasks may be changing in response to these policy shifts (Lochmiller & Mancinelli, 2019; Sun, 2012).

Although several studies suggest that the time allocation of education personnel is important for student learning and other important school outcomes like school climate, most of the research has been focused on teachers and principals (Berliner, 1990; Grissom et al., 2013; Horng et al., 2010; May et al., 2012; Stallings, 1980). This emphasis on teachers and principals is understandable because their behaviors have more direct connections to student learning than the behaviors of APs. But recent research on distributed leadership and other forms of shared leadership in schools suggests that the most effective principals share the work of leadership with

several school stakeholders (Heck & Hallinger, 2010; Spillane et al., 2007). Moreover, research on the theory of leadership substitutes suggests that leadership may be more appropriately conceptualized as a set of leadership duties that can be fulfilled by the principal or other organizational resources (Firestone, 1996; Freeston, 1987; Pitner, 1986). Despite research that suggests school leadership is not solely a function of the principal, the research on APs' duties is underdeveloped.

II.1.3 How APs allocate their time

The research on APs' work duties and tasks finds that they are most frequently asked to manage student discipline, manage relationships with education personnel, and, more recently, engage in instructional leadership (Austin & Brown, 1970; Glanz, 1994; Hausman et al., 2002; Pellicer et al., 1988; Sun, 2012). The managerial duties of APs are often described as the "Bs" of school administration, and this list of "Bs" include buses, behinds, books, buildings, bells, and balls (Good, 2008; L. J. Searby et al., 2015; Zellner et al., 2002). Of all the duties APs fulfill, managing student behavior has received outsized attention in the literature (Conley et al., 2007; Glanz, 1994; Hausman et al., 2002; Sun, 2012; Williams et al., 2020).

Many of these previous studies have described APs work as being filled with too few instructional leadership tasks, and don't find much value in the role APs perform as disciplinarians (Glanz, 1994; Koru, 1993; Marshall & Hooley, 2006). There is some evidence to suggest that this characterization of APs work may not be reflective of how APs perceive of their work in more recent years. A few studies that have specifically studied APs of color find evidence to suggest that the Disciplinarian role of APs can be valuable especially if it leads to more equitable implementation of exclusionary discipline practices (Clayton & Goodwin, 2015;

Williams et al., 2020). However, there are also studies to suggest that particularly Black APs are expected to be "race specialists" which can add stress on Black APs who already face many unique challenges in navigating school leadership (Moore, 2013). Additionally, some recent research suggests that the structure of APs' work is shifting to include more instructional leadership tasks (Allen & Weaver, 2014; Neumerski et al., 2018; Petrides et al., 2014; Sun, 2012). Although recent research suggests that APs' work may be changing, the general impression from the extant literature is that APs are assigned duties that are tedious or difficult for principals (Houchens et al., 2018; McClellan & Casey, 2015; Mertz, 2006; Militello et al., 2015; Munoz & Barber, 2011).

II.1.4 What factors are related to how APs allocate their time

Although relative to the rest of the literature on APs there are more studies on the duties of APs, few of these studies address specifically what factors shape the time APs spend on different tasks. Some evidence suggests that the personal characteristics of APs like their gender, race, age, and experience level are related to differences in how they allocate their time (Hausman et al., 2002; McClellan & Casey, 2015; Moore, 2013; Walker & Kwan, 2009). Hausman et al. (2002) find that female APs tend to spend more time on instructional leadership, professional development, and personnel management tasks. Moore (2013) finds that the race of APs shaped what expectations principals, schools, and districts had for APs. Specifically, Black APs are expected to address issues related to race in the school like the discipline of Black students. In addition to personal characteristics, research suggests that the school context can also shape the work of APs (Morgan, 2018; Walker & Kwan, 2009; Williams et al., 2020). Walker and Kwan (2009) do not find any differences in the roles of APs based on the characteristics of schools in

Hong Kong, but Morgan (2018) finds that APs in schools with more FRPL and non-white students tend to spend less time on engaging in family and community relations. Williams et al. (2020) suggest that APs in urban schools may be inclined to spend more time individualizing disciplinary approaches when working with Black students.

There are several other studies that address topics closely related to the duties of APs and the contextual factors that shape them (Munoz & Barber, 2011; Oliver, 2003; Petrides et al., 2014; L. Searby et al., 2017; Sun & Shoho, 2017), but these studies do not specifically address the duties of APs. Most of these studies examine APs' preferences for different roles and duties. Most of these studies find evidence to suggest that APs prefer to do work that is related to the instructional program of schools and involves working directly with teachers and students (Metzker, 2003; Munoz & Barber, 2011; Oliver, 2003; Walker & Kwan, 2009). Although the preferences of APs for work may have some influence on what roles they fulfill in a school, several of these studies point out that APs' preferences likely play a relatively small role in determining what they do. These studies suggest that principals are the single most important factor in how APs allocate their time (Conley et al., 2007; Houchens et al., 2018; Mertz, 2006; Militello et al., 2015; Weller & Weller, 2002). However, there are no studies that I am aware of that explicitly examine how characteristics of principals or how principals allocate their time is related to the duties of APs.

II.1.5 Leadership substitutes and APs' time use

Since prior research suggests that the work of principals is influenced by leadership substitutes, the presence of leadership substitutes may also shape the work of APs (Freeston, 1987; Pitner, 1986). Following the seminal work of Pitner (1986), which applies the theory of leadership

substitutes to schools, a few studies have examined how leadership substitutes could make certain leadership roles obsolete (Firestone, 1996; Hausman & Goldring, 2001; Podsakoff & MacKenzie, 1997). The theory of leadership substitutes has two key components: first, it suggests that leadership is best described as a series of leadership functions rather than the people in formal leadership roles, and it suggests there are characteristics of organizations that can replace or render ineffective some of the functions of leaders (Kerr & Jermier, 1978). Kerr and Jermier (1978) argue that there are potentially twelve characteristics of organizations that can potentially act as substitutes for leadership: experience-training, professional orientation, indifference to rewards, task clarity, task provided feedback, intrinsically satisfying tasks, formalization, rule inflexibility, active advisory staff, cohesive work groups, low leader position power, and spatial distance between superiors and subordinates. Pitner's (1986) study found that several of these leadership substitutes often exist in school settings. The notion of leadership substitutes is especially important to considering how school leaders allocate their time because it provides a potential explanation for why the time leaders spend on tasks may vary across schools. One of the most relevant leadership substitutes to this study is active advisory staff. Since many schools have support staff like counselors, office staff, and teacher coaches who may fulfill some of the leadership roles in schools, APs and principals may not need to spend time on the duties fulfilled by support staff. Therefore, this study explicitly examines whether there is a relationship between the number of support staff in a school, a potentially important leadership substitute, and how APs allocate their time.

II.1.6 Contributions

This study investigates several unresolved issues related to how education personnel allocate their time. First, this study extends the research on teachers' and principals' time use by examining how APs allocate their time. APs' allocation of time is an important subject for research because their work has the potential to shape the conditions in schools that enable their effective operation. Additionally, no studies in the US have systematically examined how personal and school contextual factors may constrain and shape the work of APs. A few studies have looked at a handful of characteristics that have been found to shape the work of teachers and principals, but none of the studies other than Kwan and Walker's (2009) study have explicitly tested the relative importance of these factors in explaining variation in how APs allocate their time. These studies are limited to examining the demographic characteristics of students and do not examine any other school factors. This study will extend these findings by exploring how leadership substitutes are related to APs' work. Moreover, this study will be the first to examine how characteristics and time allocation of principals are related to how APs allocate their time using quantitative methods. The field needs research that explores the relationship between the duties of APs and principals' characteristics and behaviors because principals are arguably the most important factor in determining APs' duties.

II.2 The Tennessee Contexts and APs' Time Allocation

Since this study examines differences in APs' time allocation across school contexts, it is important to describe the characteristics of schools in Tennessee with APs. Many of the differences between schools with and without APs were described in Chapter 1. In Tennessee, APs are more likely to work in high schools and middle schools, and they are more likely to

work in larger schools. Schools with APs are more likely to be in suburbs and towns, and they are less likely to be in rural communities. APs are more likely to work in schools with a slightly smaller proportion of students who are FRPL eligible and have individualized education plans (IEP). Among schools with an AP, schools with multiple APs are more likely to be high schools and are more likely to be in urban and suburban schools than schools with just one AP. Schools with multiple APs also have a smaller proportion of FRPL and IEP students than schools with just one AP. Lastly, schools with multiple APs tend to be much larger than schools with just one AP.

These differences in school contexts are consistent with the pattern that would be expected based on how AP positions are allocated in Tennessee. In Tennessee, funds for APs are allocated to schools based on the average daily membership of schools. These funds are given to districts to be allocated to schools, and districts may prescribe the number of APs in a school based on their own priorities or give the decision-making power to principals. The policies governing the allocation of AP positions are especially relevant to this study because the within district similarities in AP allocation point to the possibility that there may be within district similarities in AP time allocation. In other words, APs' time allocation patterns are likely correlated within school districts. For example, a superintendent may have an expectation that APs allocate more of their time on working directly with students. Then the percent of time allocated to working with students is likely to be correlated between APs in the same district. Standard error estimates will be biased downwards if they do not account for the clustering in the data. Therefore, the inferences made in this chapter are based on standard error estimates obtained using the cluster robust standard error formula to account for clustering at the school district level.

II.3 Data and Measures

The data for this study comes from the Tennessee Department of Education in the school years that end in 2015-2018. The first data source is the Tennessee Educator Surveys (TES), which surveys teachers and administrators annually. The TES asks administrators about several aspects of their work, but in each year, there are questions that ask school leaders about what they do in their roles. From 2015 to 2017, the TES asked administrators, including APs, about how many hours they spend on various responsibilities. The answer choices for this question were none, one hour or less, one to three hours, three to five hours, five to 10 hours, and more than 10 hours. In 2018 and 2019, the TES asked about a similar set of responsibilities, but asked school leaders to assign a percentage value to the amount of time they spent on different duties. The response counts of APs on the TES surveys and the overall response rates on the TES are reported by year in Appendix Table II.A1. A comparison of the characteristics of APs who responded to the survey to APs who did not respond to the survey is presented on Table II.1.

In addition to changes in the response format to the questions about time allocation, the responsibilities covered in each year of the TES changed slightly. The list of tasks and the years they are included in the TES are presented in Appendix Table II.A2. Although the list of items is not comprehensive, it represents a broad set of leadership tasks. The responsibilities asked about on the TES cover a range of tasks that have been identified in research as central components of the work of school leaders (E. B. Goldring et al., 2008; Grissom et al., 2013; Hausman et al., 2002; Horng et al., 2010). Table A II.2 shows that while APs are asked about instructional leadership-related tasks in each year, the same tasks are not covered in every year. Therefore, I collapse the tasks labeled as instructional tasks in Table A II.2 into a single index to allow for consistent analyses across years. No consistent definition of instructional leadership exists across

the literature, but most researchers agree that instructional leadership consists of the leadership behaviors that leaders engage in to shape school instruction (Neumerski, 2013). I label tasks as instructional leadership tasks if they appear in prior studies as instructional leadership tasks, but I also conduct a few statistical analyses to ensure the items in the constructed index fit together. First, the average percent of time APs spend on the created index of instructional leadership is relatively stable across years. I present this pattern and the percentage of AP time spent on each leadership role graphically in Figure II.1. Second, the items in the instructional leadership index have an average reliability of .38. The reliability coefficients of the index suggest that the measure is moderately reliable, and considering that the tasks are distinct, a school leader's increased time spent on one instructional leadership task may not result in an equivalent increase in another instructional leadership task. The reliability coefficients for each item and the overall index of instructional leadership across years are presented in TableII.A4.

The second data source I draw from in this study is the data from the Tennessee Educator Acceleration Model (TEAM), the multiple measure evaluation system for educators in Tennessee. The TEAM evaluation system was introduced in the 2011-2012 school year and evaluates teachers and administrators based on a combination of student test scores and subjective ratings by supervisors. All administrators (both principals and APs) are rated by supervisors using a rating rubric that mirrors the domains covered by the Tennessee Instructional Leadership Standards (TILS)[7]. The TILS are approved by the state board and grounded in the Professional Standards for Educational Leaders[8], formerly the ISLLC standards. Principals are

[7] The current version of the TILS standards can be found here: https://www.tn.gov/content/dam/tn/stateboardofeducation/documents/policies/5000/5.106%20Tennessee%20Instructional%20Leadership%20Standards%20Policy%207-27-18.pdf

[8] The PSEL standards can be found here: http://npbea.org/wp-content/uploads/2017/06/Professional-Standards-for-Educational-Leaders_2015.pdf

evaluated by district or state personnel who are tasked with supervising principals, which could be a superintendent or another central office supervisor. According to State Board policy 5.201 on educator evaluation, districts can select any state approved frameworks to evaluate school leaders. Currently, there is only one state-approved alternative evaluation framework for school administrators, and most districts use the rubric developed by TDOE. There are four domains in the current version of the administrator TEAM rubric designed by TDOE: instructional leadership for continuous improvement, culture for teaching and learning, professional learning and growth, and resource management[9]. In each domain there are multiple indicators, and AP supervisors are asked to rate APs on a five-point scale for each of the indicators. These scores are averaged to produce a score for each domain and an overall score. This study focuses on the overall evaluation scores for the primary analyses predicting APs' time use because all districts must provide an overall score and prior research suggests that the TEAM scores on individual indicators are derived from a single underlying factor (Grissom et al., 2018).

The third data source is the administrative staff files that provide demographic characteristics of personnel including gender, race[10], age, number of years in Tennessee's education system, and their position in a given school year. Experience as a teacher, AP, principal, and years in the same position have been created from the administrative staff files. Since the administrative files only go back to 2002, there is systematic undercounting of experience at the higher range of personnel experience. This is less of an issue in the years of this study because most school leaders in the sample (approximately 97 percent) have their entire careers as school leaders observed in the data. The final data source is the student administrative

[9] The TEAM rubric can be found at https://team-tn.org/wp-content/uploads/2013/08/TEAM-Administrator-Rubric.pdf

[10] There are very few APs who do not identify either as white or Black, so I group APs from all other racial identifications together in the category of "Other."

files, which have students' gender, race, FRPL, and IEP status. These student-level data are used to make school-level averages. All data are obtained from the Tennessee Education Research Alliance at Vanderbilt University.

II.3.1 Measures capturing the time allocation of APs

To descriptively explore how APs time is allocated, I create variables that represent different approaches to identifying the time APs spend on various leadership tasks. The first approach re-codes the responses for APs' time allocation in 2015-2017 to turn them into percentages. This makes the responses from 2015-2017 comparable to those in 2018 and 2019. To make these transformed variables, I take the midpoints of the response categories and assign that midpoint in place of the response category indicator. For example, the midpoint of the response one to three hours is two hours, so all APs who chose one to three hours would be recoded as a two. Those APs who chose more than 10 hours will be recoded as 15 hours which is 1.5 times 10, the largest number of hours that APs can report on the survey. I then add up the total number of hours across all the available tasks. I create percentages of time spent on different tasks by then dividing the hours spent on a specific task by the total number of hours spent on all tasks. The potential danger with using this sort of recoding scheme is that measurement error is being introduced as part of the recoding, particularly for the recoded responses of 10 hours or more to 15 hours. Therefore, as a sensitivity check I run all analyses using two other recodes of the largest categories as 10 and 12.5, which uses the lowest possible value in the top category and adds half the range from the previous category respectively. Although the point estimates change slightly, all the substantive findings of this study are consistent across these different recodes of APs' time spent on different duties.

There are eight different types of duties that the TES asks school leaders about: discipline, instructional leadership, administrative, parent communication, meetings with central office, working directly with students, supervisory, and other duties. All of these categories are addressed every year except for the other duties category, which was only asked in the 2018 and 2019. As mentioned earlier, the specific duties asked about on each year of the TES are presented in Appendix Table II.A2. Most of these duties are intuitive, but I will define each type of duty as described on the TES for the sake of clarity. The discipline category refers to the time that school leaders spend on managing student discipline. The instructional leadership index created using all the categories in instructional leadership refers to an assortment of duties related to the instructional program of schools. This category includes observing teachers, providing feedback, planning instruction, coaching teachers, modelling lessons, evaluating teachers, and other instructional leadership tasks. The most divergent year is 2019, which asks only about evaluating teachers and other instructional leadership tasks. The alpha coefficient for the index created from these two items is the weakest of all the years. However, year-by-year analyses confirm the substantive findings of this study despite these differences in the instructional leadership category. Administrative duties primarily refers to paperwork that needs to be completed for compliance purposes but it can also include human resources and managing facilities. Parent communications is a category that refers to what school leaders do to communicate with parents but also community members. This category on the TES includes both of these stakeholders in the item each year. Meetings with central office refers to meetings either initiated by or with central office personnel. Working directly with students may refer to duties like counseling students, but may also include meetings that involve students and another school

community member. Supervisory duties include supervising school activities but also monitoring the lunch room and hallways.

The second type of variable that captures the types of tasks APs spend time on is created using a cluster analysis similar to the procedure described in Goldring et al.'s (2008) study of principal time allocation. Cluster analysis is a multivariate data reduction technique that uses an algorithm to group observations to minimize the within-group variation and maximize the across-group variation. The cluster analysis approach used in Goldring et al. (2008), and the most commonly used cluster analysis approach, is the k-means method. In this procedure the algorithm searches around the data for groupings of observations that maximize the distance between a pre-specified *k* number of group means along the variables included in the estimation. I conducted this analysis using the cluster command in Stata 15.[11] Although this procedure produces clusters of observations, there are no statistical tests that explicitly inform users of the optimal number of clusters. Makles (2012) describes a series of graphical and statistical tests that can be employed to inform how users can optimize the number of clusters in a cluster analysis using Stata's cluster command. Makles (2012) suggests creating clustering solutions ranging from one to 20 clusters and capturing the within sum of squares (WSS), the log of the WSS, the η^2, and the proportional reduction of error for each of these solutions. Then these statistics should be plotted to look for "elbows" in the curves that indicate the number of clusters that should be used in a solution. The results from this procedure conducted on the sample of APs are presented in Figure II.2. Figure II.2 suggests that a five-cluster solution is the most appropriate for the data in this study. Table II.2 describes the patterns in time use for APs in each of the five clusters, and Figure II.3 presents these results graphically. Based on the time allocation profiles of APs in

[11] StataCorp. 2017. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.

each of the five clusters, I give more intuitive names to the cluster profiles. The names of the five AP time allocation profiles are Administrators, Supervisors, Generalists, Disciplinarians, and Instructional Leaders.

II.3.2 Measures of support staff in schools

Appendix Table II.A3 presents a list of position titles of support staff commonly found in schools during the 2015–2018 school years. I define support staff as all personnel in schools who are not teachers, APs, or principals. Although all the support staff in a school may not act as advisory staff, this study explores the relationship between the number of support staff as reported in the data and APs' time allocation because Pitner (1986) finds evidence to suggest that leadership substitutes shape principal behavior. The research on support staff is relatively thin, and few studies have examined how school support staff influence the behavior of school leaders (Blatchford et al., 2006; Whitehorn, 2010). However, Figure 1 suggests APs spend substantial time on instructional leadership, discipline, administrative tasks, and working directly with students. Appendix Table II.A3 includes position titles for support staff who may have a direct influence on the amount of time APs allocate to these top duties. For example, having a school counselor in the building may reduce the amount of time APs spend on school discipline because some of the work may be shared with school counselors. To capture the number of support staff in schools who may specifically address the leadership tasks asked about on the TES, I categorize the positions of support staff in the study data based on categories of state data found in the LEA universe survey.[12] These categories are counselors, instructional coordinators and

[12] The LEA universe survey and the data dictionaries can be found at https://nces.ed.gov/ccd/pubagency.asp

supervisors of staff (instructional), librarians/media (materials), administrative, and student support.

II.4 Methods

II.4.1 Descriptive Analyses

The analyses for this study begins with a descriptive analysis of APs' time use. I first conduct *t*-tests to examine differences in APs' personal and school characteristics across different time use profiles. Next, I compare APs' time use across grade levels and locale types. These descriptive analyses show how APs' personal and school characteristics differ across time use profiles, and how APs' time use differs across school contexts. The limitation of these analyses is that they do not examine the relative importance of these factors. To examine the relative importance of different factors in how APs allocate their time, I conduct the regression analyses described in the next section.

II.4.2 Regression Analyses

The analyses for this study are modeled after the analyses in Grissom et al. (2015), which regresses percent of time allocated to duties on personal and school characteristics. The outcome measures used in this study are the percent of time spent on duties and the likelihood of being categorized into one of the time allocation profiles created from the cluster analysis. All regression models take the form described in Equation 1.

$$Duty_{kist} = \alpha + APChars'_{ist}\beta' + SchoolChars'_{st}\theta' + Support'_{st}\mathrm{X}' \qquad (2)$$
$$+Principal'_{st}\Pi' + OthersDuty'_{k'st}\mathrm{M}' + \tau_t + \epsilon_{ist}$$

$Duty_{kist}$ represents the different measures of time allocation *k* by AP *i* in school *s*, district *d*, and year *t*. $Duty_{kist}$ includes the percent of time on specific duties and time use profiles. The models that have time use profiles as outcomes are linear probability models. $APChars'_{ist}$ is a vector of APs' demographic characteristics (gender, race, age) and APs' experience (years as a teacher, years as an AP, years in the same position, and educational attainment). $SchoolChars'_{\mathrm{st}}$ is a vector of school characteristics (proportion Black students, proportion Latinx students, proportion other race students, proportion FRPL eligible students, proportion of students with IEPs, school enrollment in 100s, grade level, locale type, and the number of APs in the school). For some of the models I include $Support'_{st}$,which represents the counts of support staff in schools. $Principal'_{st}$ represents the demographic characteristics, evaluation ratings, and experience levels of principals, and $OthersDuty'_{k'st}$ is a vector of the time principals and other APs in the school allocate to leadership duties, and τ_t represents year fixed effects. The results from these regression analyses should reveal the relative importance of factors that are related to APs' time allocation.

II.5 Findings

II.5.1 Descriptive analyses

The descriptive analyses I conduct to answer the first research question are presented in Tables II.2-II.6 and Figures II.3-II.4. Figure II.3 presents the time allocation of principals and APs side by side to facilitate comparisons. Descriptively, principals tend to allocate more time to instructional duties and less time to disciplinary duties than APs. Principals allocate about 36 percent of their time to instructional duties and 13 percent of their time to discipline. APs allocate about 31 percent of their time to instructional duties and 23 percent of their time to

discipline. The differences in time allocated to instruction and discipline are the most apparent differences in the way principals and APs allocate their time. Principals allocate about 4 percent more of their time to administrative duties than APs, and principals allocate a little less time meeting with students and supervising activities than APs. This figure suggests that APs tend to allocate their time differently, especially as it relates to the three primary leadership duties of instruction, discipline, and administration.

Table II.2 reports the percent of time APs spend on duties across the different time allocation profiles produced from the cluster analysis. The information from Table II.2 is displayed graphically in Figure II.4. All of the comparisons in this table are to the first column, the Administrators. The first observation from this table is that there are more Generalists than any other type of AP time allocation profile. The profile that describes the second most APs is the Instructional Leader profile. A second finding from this analysis is that although APs from specific profiles may allocate more time to some duties, all APs tend to spend some time on every type of duty. Additionally, prior research suggests that APs spend a substantial amount of their time on discipline, but this analysis suggests that discipline occupies about 20 percent or less of an APs' time for four of the five profiles. It is also noticeable that for a large number of APs, the largest percentage of their time is spent on instructional leadership tasks. For both the Generalists and Instructional Leader profiles instructional leadership is the duty that takes up the most time.

Table II.3 conducts *t*-tests of the differences in APs' personal and school characteristics across time profiles. All comparisons in this table are also made with the profile in the first column, Administrators. This analysis suggests that Instructional Leaders are more likely to be female, and both Supervisors and Disciplinarians are less likely to be female than the other

profiles. A larger percentage of Generalists, Disciplinarians, and Instructional Leaders are likely to be Black than Administrators. Supervisors tend to have the most years of experience as APs and tend to be the oldest. Disciplinarians have significantly lower evaluation ratings than all other AP profiles. Instructional Leaders ted to work in schools with the most students of color. Instructional Leaders also tend to be in the smallest schools and are more likely to work in elementary schools than all other grade levels. Instructional Leaders are the most likely to be in urban schools as compared to all other time allocation profiles. Supervisors tend to be in larger schools and in high schools. Supervisors are the most likely to be in rural schools as compared to all other profiles. Administrators work in schools with significantly fewer FRPL eligible students and are most likely to work in suburban schools as compared to all other profiles. Disciplinarians tend to work in schools with the most FRPL eligible students and are more likely to be found in middle schools compared to all other profiles. Disciplinarians tend to be the least likely to be in urban schools and are the most likely to work in schools in towns. Disciplinarians also work in schools that have significantly lower achievement scores than any other profile. There are no significant differences in principals' evaluation ratings across APs' time use profiles. There are differences in the number of support staff across profiles, and Disciplinarians tend to have more support staff and Instructional Leaders tend to have less support staff.

Table II.4 compares time use of APs across different grade levels. All comparisons in this table are to elementary school APs. The results from this descriptive analysis suggest that high school APs are the most likely to be Administrators and Supervisors, and they are the least likely to be Instructional Leaders or Generalists. Consistent with these findings, APs in high schools are more likely to say that administrative tasks and supervisory tasks make up the largest percentage of their time. High school APs are also the most likely to report that working directly

with students is the task that takes up the largest percentage of their time. Middle school APs are the most likely to fit the Disciplinarian profile and are the most likely to report that student discipline takes up the largest percentage of their time. Elementary school APs are the most likely to be Instructional Leaders or Generalists and are the most likely to report that instructional duties take up the largest percentage of their time. Approximately 73 percent or all elementary school APs report that instructional duties take up the most time in their work.

Table II.5 reports the results from comparing time use of APs across schools in different locales. All comparisons in this table are to urban school APs. Urban school APs are the most likely to be Generalists and Instructional Leaders and are the least likely to be Supervisors. APs in suburban schools are more likely to be Administrators than APs in any other locale. APs of schools in towns are the most likely to be Disciplinarians, and APs of rural schools are the most likely to be Supervisors. APs of urban schools are the most likely to say that instructional duties take up the largest percentage of their time. APs of schools in towns are the most likely to report that discipline takes up the most time, and rural APs are the most likely to say that supervisory duties take up the largest percentage of their time. Across all locales, instructional leadership is the duty that APs are the most likely to report as the largest percentage of their time.

II.5.2 Regression Analyses

To answer the second research question, I regress APs' percentage of time allocated to duties and time use profiles on personal and school characteristics. All analyses include year fixed effects and cluster standard errors by school district. In table II.6 the outcome measure is the percentage of time APs spend on different duties. Starting with the personal characteristics of APs, these analyses suggest female APs tend to allocate significantly less time on discipline and supervisory

duties. Female APs spend about 3.3 percentage points less time on student discipline and about 2.4 percentage points less time on supervisory duties than male APs. Lastly, female APs allocate about 1 percentage point more of their time on administrative duties as compared to male APs.

Black APs allocate significantly less of their time on administrative duties spending about 3.6 percentage points less of their time on administrative duties than white APs. Black APs spend about 1.2 percentage points more of their time on parent and community relations and about 1.1 percentage points more of their time meeting with students than white APs. Older APs spend slightly less time on administrative duties. A one-year increase in age is associated with a .1 percentage point decrease in time allocated to administrative duties. More experience as an AP and as a teacher is associated with more time on supervisory duties, and a one-year increase in experience as an AP is associated with about a .2 percentage point increase in the amount of time spent on supervisory duties. Lastly, APs with more teacher experience tend to allocate more time to discipline and supervisory duties, and they spend less time meeting with students. These results suggest that holding all other covariates constant, an APs' gender, race, age, and experience is related to how they allocate their time. Race and gender appear to be especially important factors in determining how APs allocate their time. However, experience and age variables, while significant, only differentiate the time allocation of APs at the tails of the distribution of those variables.

Table II.6 also presents the relationships between school characteristics and APs' time allocation. It is important to note, when reading this table, that coefficients on the proportion of students from a particular group represent a change from no students in that group to all the students being from that group. So, the coefficient on proportion IEPs represents the change in time allocated to a duty associated with a change from no students with an IEP to all students

having an IEP. APs of schools with more Black students tend to spend more time on instructional tasks, and they tend to spend less time on parent and community relations and supervising students. An AP in a school with all Black students spends about 6.7 percentage points more of their time on instructional leadership than an AP in a school with no Black students. Few schools either have all Black students or no Black students, so a more plausible difference may be between an AP at a school at the 10th percentile of Black students (about 1.6 percent of students identify as Black) compared to APs at a school in the 90th percentile of Black students (about 81.1 percent of students identify as Black). Assuming that the difference in time allocation is distributed evenly across the proportion of Black students, the difference for an AP at the 90th percentile as compared to an AP at the 10th percentile is about 5.3 percentage points and is equivalent to about .30 SD in AP time allocation on instructional duties. Similar patterns exist for schools with larger proportion of Latinx students, but there are proportionally fewer Latinx students in Tennessee schools than Black and white students.

Schools with higher proportions of FRPL eligible students have APs who tend to spend more time on managing discipline (about 7.8 percentage points associated with a change from no FRPL eligible students to all FRPL eligible students) and less time on instructional duties and less time on administrative duties (about 3.7 and 3.5 percentage points less respectively). APs in schools that have more students with IEPs tend to spend less time on supervisory duties, and in a theoretical school with 100 percent of students with an IEP, the AP would spend approximately 17.9 percentage points less time on supervisory duties than APs in a theoretical school with no students with an IEP. Although this difference seems very large, no schools have student populations that all have IEPs or none have IEPs. The difference in the proportion of students with IEPs between schools at the 90th percentile of students with IEPs and the bottom 10

percentile of schools is about .126. The predicted difference in APs' time allocation on supervisory duties between the 90th and 10th percentile schools' APs is about 2.3 percentage points. APs at schools with more students with IEPs also tend to spend more time on instructional duties and discipline, but they spend significantly less time on administrative duties.

As found in the descriptive analyses, APs in middle and high schools spend more time on student discipline, more time on supervisory duties, and less time on instructional leadership. APs in middle schools tend to spend less time on administrative duties, and APs in high schools spend less time on parent communication and more time on meeting with students than elementary school students. APs in middle schools spend about 6.4 percentage points less time on instructional duties and 8.3 percentage points more of their time on student discipline compared to elementary school APs. There are relatively fewer significant differences in AP time allocation across locale types. APs of schools in towns allocate more time on discipline (about 2.7 percent) than APs of urban schools while APs in rural schools tend to spend more time on supervising activities (about .8 percent) as compared to APs in urban schools. Lastly, APs in schools with higher levels of achievement tend to spend more time on instructional duties and spend less time on discipline and administrative duties. The AP of a school one stand derivation higher on student achievement spends about 1.6 percentage points more time on instructional duties, about 1.3 percentage points less on discipline, and about .8 percentage points less on administrative duties. These results do not necessarily suggest that if an AP spends more time on instruction and less time on discipline that student achievement will improve. It does suggest that there is a pattern of behavior that fits with the long-held theory that instructional leadership of APs can contribute to student learning.

Table II.7 reports the results from the linear probability models of having a particular time allocation profile on personal and school characteristics.[13] Gender and race are predictive of time allocation profiles in addition to percentage of time spent on specific activities. Female APs are about 8.6 percent less likely to be Supervisors and 4.4 percent less likely to be Disciplinarians and female APs. However, female APs are about 8.5 percent more likely to be Instructional Leaders as compared to male APs. Black APs are about 7.5 percent less likely to be Administrators than white APs. APs in schools with more Black students are more likely to be Instructional Leaders, but they are less likely to be Supervisors. APs in schools with more Latinx students follow a similar pattern to schools with more Black students, but APs in schools with more Latinx students are also much less likely to be Disciplinarians. APs in schools with more students with IEPS are less likely to be Supervisors or Administrators, and they are more likely to be Generalists or Disciplinarians. When compared to elementary school APs, middle school APs are about 16.4 percent more likely to be Administrators and 1.44 percent less likely to be an Instructional Leaders than elementary school APs. High school APs are about 11.1 percent less likely to be categorized as an instructional leader and about 6.5 percent less likely to be Generalists as compared to elementary school APs. High school APs are also about 7.6 percent more likely to be categorized as Disciplinarians and about 8.6 percent more likely to be categorized as a Supervisors compared to elementary school APs. APs of schools in towns are about 3.5 percent likely to be categorized as Supervisors compared to urban school APs, and APs of rural schools are about 3 percent more likely to be Supervisors as compared to urban school APs.

[13] I also conducted these same analyses as multinomial logit models and found that the results lead to substantively similar inferences.

Tables II.6 and II.7 seem to suggest that the personal characteristics of APs and the characteristics of the schools they work in are related to how they allocate their time. These tables do not explain what the underlying reason for these relationships is, but they suggest that either the characteristics of APs and their schools or some unobserved connection to those factors is associated with how APs allocate their time. Among the school factors that seem to have the most consistent relationships with APs' time allocation are grade level, the demographic characteristics of the student population, and the achievement level of the school. APs in higher grades are more likely to allocate more time towards student discipline and less time to instructional duties. APs in schools with more students of color are more likely to allocate more time to instructional duties and less likely to allocate time to other duties while APs in schools with more FRPL eligible students are more likely to allocate their time to discipline. Lastly, APs in schools with higher achievement levels tend to spend more time on instruction and less time on discipline or administration.

To answer the third research question, I estimated regressions of the percentage of time use on duties and the likelihood of being categorized as a specific time allocation profile on the number of support staff in a school. This question addresses how one potential leadership substitute, active advisory staff, is related to how APs allocate their time. These results are reported in Tables II.8 and II.9. *Panel A* of these two tables presents the estimated relationship between the total number of support staff and the time allocation outcome listed at the top of the column, and *Panel B* presents the estimated relationships between the number of specific support staff positions and time allocation outcomes. Table II.8 suggests that APs in schools with more support staff tend to spend more time on instructional leadership and less time on supervisory duties. *Panel A* of Table II.8 suggests that one additional support staff in a school is associated

with APs spending .17 percent more of their time on instructional leadership and .09 percentage points less time on supervisory duties. The results in Table II.8 *Panel B* suggest that much of the relationship between the total number of support staff and APs' time use can be explained by the relationships with the number of instructional and materials support staff. This pattern seems to fit with the fact that instructional and materials support staff are some of the most common support staff in schools. Having one additional instructional support staff is associated with .17 percentage points more time spent on instructional duties and .13 percent less time on supervisory duties. Having one additional materials support staff is associated with APs spending .93 percent more time on instructional duties. Lastly, having one additional student support staff is associated with .31 percent less time on meeting directly with students, and having one additional counselor is associated with allocating .13 percent more time on meetings with central office.

Table II.9 suggest that the total number of support staff in a school only has a significant relationship with the likelihood of being categorized as an Instructional Leader. Having one additional support staff in a school is associated with a .4 percent increase in the likelihood of being categorized as an Instructional Leader. The specific support staff positions that are significant predictors are instructional and student support staff. Having one additional instructional support staff is associated with a .4 percent decrease in the likelihood of being a Supervisor. Having one additional student support staff in a school is associated with a 1.6 percent decrease in the likelihood of being categorized as a Generalist and a 1.3 percent increase in the likelihood of being categorized as an Instructional Leader. The results from Table II.8 and II.9 suggest that support staff do not act as substitutes for AP leadership as described by Pitner (1986). For example, we would expect that having additional instructional support staff should

result in APs spending less of their time on instructional duties if instructional support staff act as substitutes for APs' leadership on instructional duties. Instead, APs appear to spend more time on instructional duties when they have more instructional support staff. Support staff also seem to have a substantively small relationship with the time allocation of APs. Although there are some significant relationships between having an additional support staff and APs' time allocation, there are only substantively small changes in Aps' time allocation associated with differences in the number of support staff.

Tables II.10 and II.11 address the fourth research question about the relationship between principals' characteristics and how APs allocate their time. Table II.10 suggests that there are no significant differences in APs' time allocation across the race and gender of their principals. Principals' experience as a principal and specific experience as the principal of the same school also have a significant relationship with time allocation of APs. Having a principal with one additional year of experience as a principal is associated with an AP spending about .12 percent more time on supervisory duties. Having a principal with more experience as the principal of the same school is associated with APs spending more time on meeting directly with students (about .18 percent) and an almost equal percent of their time less on supervisory duties. There is also a significant but substantively small relationship between additional years of experience as the principal of the same school and APs having to allocate more time to meetings with central office personnel. The characteristic of principals who have the largest relationship with APs' time allocation is the principals' evaluation rating. Principals with higher evaluation ratings have APs who allocate more time to instruction and administration, but those APs also tend to spend less time on discipline and supervisory duties. A one standard deviation difference in evaluation ratings of principals is associated with a 2.7 percentage point increase in APs' time on

instructional duties and a 2.5 percentage point increase on administrative duties. A one standard deviation difference in principals' evaluation ratings is also related to a 3.6 percentage points decrease in time allocated to discipline and a 1.1 percentage points decrease in time allocated to supervision.

Table II.11 presents the results from regressing AP time profiles on principals' characteristics. This table suggests that the evaluation rating of the principal is the only characteristic of principals that is significantly related to the time use profiles of APs. APs with higher rated principals are more likely to be Instructional Leaders or Administrators and less likely to be Disciplinarians or Supervisors. A one standard deviation increase in principal evaluation rating is associated with a 7.9 percent increase in the likelihood an AP is an Instructional Leader and a 6.1 percent increase in the likelihood an AP is an Administrator. A one standard deviation increase in principal evaluation rating is also associated with a 7.8 percent decrease in the likelihood an AP is a Disciplinarian and a 4.7 percent decrease in the likelihood an AP is a Supervisor.

Tables II.10 and II.11 suggest that principals do matter for how APs allocate their time. Both a principal's experience as a principal and specific experience as the principal of the same school are related to how APs allocate their time. More experience as a principal, holding experience as a principal in the same school constant, is associated with more time allocated to supervisory duties. However, more experience as the principal of the same school, holding total experience as a principal constant, is associated with less time allocated to supervisory duties. It seems overall that the two experience measures seem to have opposing relationships with APs' time allocation. Although experience as the principal of the same school and total experience as a principal are equivalent in the early years of a principal's career, these variables are likely to

diverge later in a principal's career. The most helpful way to think about the interpretation of these coefficients may be to think of the case of two different types of principals, one principal who has been only the principal of the same school for five years and a principal who is new to a school but has five years of experience as a principal. Holding all other regressors in the model at their means, an AP with a principal who only led the same school for five years is predicted by the analytic model to allocate 7.7 percent of their time to supervisory duties. An AP with a principal who has had five years of experience as a principal but is new to the school is predicted to allocate 7.1 percent of their time to supervisory duties. This example illustrates how the substantive differences in APs' time allocation across principal experience is small despite being significant. This difference becomes more pronounced as principal experience increases and for comparisons of APs in schools led by principals at the tails of the distribution of experience.

The more important principal factor for APs' time allocation is the principals' evaluation ratings. Principals with higher evaluation ratings have APs who allocate more time to instructional duties and administrative duties, and their APs allocate less time to disciplinary and supervisory duties. Similar to the pattern observed regarding the achievement level of schools and the time allocation of APs, this model cannot point to the causal direction between the time allocation of APs and the evaluation ratings of principals. The results are consistent with prior research that suggests instructional leadership and administrative management are important functions fulfilled by principals. It may be that successful APs, like successful principals, tend to allocate more time to instructional leadership and administration. Moreover, it may be that higher rated principals are better coaches for APs and give their APs more experience in the areas that are critical for success as a leader.

Tables II.12 and II.13 present the results of regressing an APs time allocation on the time allocation of other school leaders. Since the full model that includes both the time allocation of other APs and principals excludes APs who do not have other APs in the school, data on the time allocation of other APs, or data on principals, the results from the full model may be sensitive to sample selection bias. To check whether these results are sensitive to sample selection, I conduct the analyses on the sample of APs with data on the time allocation of other APs in their building and the sample with data on the time allocation of their principals separately as well as together. Another potential concern for these analyses is that the time other leaders allocate to duties are likely to be highly correlated with one another. This threat of multicollinearity can be seen in the pair-wise correlations of other leaders' time allocation presented in Appendix Tables II.A5-A7. This concern becomes more important as more duties are included in the model because the time allocated to any one duty is perfectly predicted by the combined time allocated to all other duties. When examining the variance inflation factors (VIF) of the regressors in a model that includes all the duties of other school leaders, I find that the almost all the duties have a VIF above 10 and the overall model has a mean VIF of 11.7. If I only include the time other leaders allocate to the three primary leadership duties of APs (discipline, instruction, and administration), none of the coefficients on duties has a VIF above five and the model has an overall VIF of 2.15.[14] Therefore, the models that examine the time APs allocate to the three primary duties of APs only include the time other leaders allocate on those same three duties. In models that examine the time APs allocate to minor duties, I include the time other leaders allocate to the minor duty and the three primary duties.

[14] The general rule of thumb for applied econometric work is that a VIF above the threshold of 5 is considered a moderate problem and above the threshold of 10 is considered a serious problem (Wooldridge, 2015)

The results presented in Table II.12 suggest that the time APs allocate to instruction, discipline, and administration have little relationship with the time other APs allocate to those duties. The only marginally significant relationship is between time spent on administrative duties and time spent on administration by other APs, but this relationship is only significant in the model that also includes principal time allocation. On the other hand, the time that principals allocate to different duties does seem to be related to the time APs allocate to their duties. APs tend to spend more time on instruction and discipline as their principals spend more time on other duties. So, a 1 percent increase in the time that principals spend on discipline is associated with a .18 percent increase in the time APs spend on instruction. Administrative duties do not seem to follow this same pattern because APs tend to spend more time on instructional duties as their principals spend more time on administrative duties, but APs do not tend to spend more time on administrative duties when their principals spend more time on discipline or instruction. This suggests that principals may think of APs as substitutes to their leadership in instructional and disciplinary duties, but principals may not think of APs as substitutes for their leadership in administration.

The analyses reported in Table II.13 find that the percent of time that other leaders allocate to duties has a slightly different relationship with the time APs allocate to minor duties. Whereas APs time allocated to primary duties generally has a positive relationship with principals' time allocated to other duties, both other APs and principals' time allocated to minor duties seem to have a positive relationship with time allocated to the same minor duty. For example, a 1 percentage point increase in the time a principal allocates to meetings with students is associated with a .07 percentage point increase in the time APs allocate to meeting with students. This pattern is true of all four minor duties and principals' time allocated to minor

duties. There is a similar pattern for the time APs allocate to a minor duty and the time other APs allocate to minor duties. A 1 percent increase in the time other APs allocate to meeting with students is associated with a .1 percent increase in the time APs allocate to meeting with students. Some of this relationship between other APs' time allocated to a minor duty and time APs allocate to the same minor duty is a function of the relationship between the time principals allocate to a minor duty and the time APs allocate to the same minor duty.

The models that include both other APs' and principals' time allocations suggest that there is substantial loss of precision because the point estimates are similar, but the standard errors are much larger on the same duty. In the models with both other APs' and principals' time allocation, the time principals allocate to instructional duties is associated with APs allocating more time to parent relations, meetings with central office, and meetings with students. Holding the time allocation of all other leaders on other duties constant, a one percent increase in principals' time allocated to instructional duties is associated with about a .12 percentage point increase in the time that APs allocate to these three different types of meetings. The relationship between principals' time and APs' time on meetings suggests that APs are assigned some of these minor duties to free up their principals' time for instructional duties. It is important to note that this relationship is only true of the stylized sample of APs who have other APs and those other APs do not experience an increase in the time they allocate to these minor duties. This result suggests that in some cases a principal with multiple APs may ask one AP to focus more time on with parents, central office, and students. There were no discernable patterns in the relationships between the time allocation of principals and other APs and the time use profiles of APs. Although the time APs allocate to duties has a significant association with the time other

APs and the principal in the building allocate to duties, those relationships are substantively small and are not likely to matter for larger patterns in time use.

II.6 Discussion

The results from these analyses extend and update many of the findings of prior research on the time allocation of education personnel. First, they suggest that APs who are often thought of as primarily addressing the "Bs" of school administration actually spend the most time on instructional duties. Although APs spend substantial portions of their time on school discipline and administrative duties, it is too simplistic to characterize their role as being defined by the time they spend on discipline and administration. APs allocate their time differently from principals, but there is substantial overlap in duties with principals. Second, the results show that APs' work is contextual just as principals' work is contextual (Goldring et al., 2008). APs' work differs substantially across grade levels and school characteristics, and these differences suggest that APs allocate their time according to the needs of their schools.

This study also suggests that support staff and the time allocation of other members of the school leadership team are related to how APs allocate their time. Although the number of support staff in a school is related to the percentage of time that APs spend on different duties, support staff do not appear to behave primarily as leadership substitutes. If support staff acted as leadership substitutes, we would expect an AP's time allocation to duties to decrease as the number of support staff in an associated area increases. For example, the number of instructional support staff in a school is positively related to the percentage of time that APs allocate to instructional duties. If support staff acted as leadership substitutes we might expect APs to actually spend less time on instructional leadership as they have more instructional support staff.

While the time allocation of other APs in the building does not seem to have much of a relationship with the time APs allocate to the three primary duties of instruction, discipline, and administration, the time APs allocate to minor duties does seem to be related to the time other APs allocate to the same minor duty. However, this relationship seems to be partially a function of the relationship between principals' time allocation and APs' time allocation.

Lastly, the findings from this study suggest that principals matter for APs' time allocation. Although the belief that principals share leadership with APs has been a long-held assumption, this study documents the characteristics and behaviors of principals that are associated with the time allocation of APs. Holding all other characteristics of schools and APs constant, principal personal characteristics like race and gender are not predictive of APs' time allocation, but the evaluation ratings and experience of principals is significantly related to the percentage of time APs allocate to duties. APs who serve under principals with higher evaluation ratings are likely to allocate more time to instructional and administrative duties while allocating less time to discipline and supervision. Interestingly, APs that work with high evaluation rating principals allocate more time to the same areas that more effective principals allocate time towards in prior literature (E. B. Goldring et al., 2008; Grissom, Loeb, et al., 2015; Horng et al., 2010). This same pattern in APs' time allocation is also found in the relationship between APs' time allocation and the achievement level of schools. APs in schools with higher levels of achievement tend to allocate more time to instruction and administration and less time to discipline. These relationships hold even when controlling for all other characteristics of schools.

II.7 Limitations

II.7.1 Threats to internal validity

A particularly salient concern for this study is the issue of measurement error associated with the time allocation variables. The first concern is that measurement error, specifically a ceiling effect, will lead to attenuation of the estimated relationships. As described in the data section, I recoded the variables into percentages from ordinal answer choices for responses from the 2015, 2016, and 2017 administrations of the TES. This operationalization of the variable will add measurement error into the variables because the answer choices in 2015 to 2017 impose a ceiling on how many hours APs can report having worked on a duty. The primary concern with the measurement error that may occur as part of the recoding of the variables is that variation in time allocation is suppressed. Since some of the factors analyzed in this study have substantively small relationships with the time allocation of APs and this measurement error attenuates the estimates, there may be some relationships that are actually significant in reality but are not detected in my analyses.

Another potential source of measurement error in the time allocation variables is that APs could only report spending time on the areas that are asked about on the TES in any given year. There is a possibility that APs' responses to some of the questions may have been different if more roles were asked about. For example, in 2015 and 2016 the TES asked APs how many hours they spent coaching teachers, but in 2017 the choices were expanded to include modelling lessons in addition to coaching teachers. If the APs in 2015 and 2016 would have reported a different amount of time spent on coaching teachers had they been prompted to report how much time they spent on modelling lessons, then there would be measurement error in the hours spent

on coaching teachers in 2015 and 2016. The most likely result of this measurement error would also be attenuation of the estimates.

Although this source of measurement error is not directly addressed by Figure II.1, it is somewhat reassuring to find evidence that the amount of time spent on the various duties appears to be relatively stable over time. This descriptive pattern suggests that the changes in the categories and response formats does not seem to dramatically change the percentage of time that APs report spending on various duties. Additionally, both potential sources of measurement error in this study are likely to result in attenuation bias if present. The practical consequence of this attenuation bias would be an inability on the part of the estimators to detect significant relationships. Despite these concerns, I find significant relationships between multiple factors and the percentage of time APs allocate to different duties even when the estimated relationships are relatively small.

A final concern may be that there are omitted variables that explain the relationship between the factors included in this study and the percentage of time APs allocate to duties. For example, this study finds a significant relationship between the gender of APs and their time allocation. There is a possibility that some unobserved factor about APs and the time they allocate to duties explains that relationship. However, this type of bias would be a much larger concern if I were attempting to make a causal claim about the gender of APs and how they allocate their time. This study only attempts to describe patterns between variables that may be relevant to how APs allocate their time and the percentage of time APs allocate to duties. I also include a robust set of covariates in this study, so whatever unobserved factors may bias the relationships estimated in this study would have to be unrelated to the covariates included in the models.

II.7.2 Generalizability

Another salient concern in this study is the response rates on the TES. I can only analyze the data for APs' who responded to the TES. This means that this study is not generalizable to the types of APs who did not respond to the survey. The comparison of APs who responded to the survey and APs who didn't respond to the survey reported in Table II.1 suggest that respondents are significantly different on observable factors from non-respondents. Since APs in the sample are systematically different from APs not included, non-respondents may allocate their time in ways that are systematically different from respondents. Therefore, this study lacks external validity to samples of APs who are different from those that responded to the survey.

Lastly, this study is set in one state over a few years. The patterns in Tennessee may not be reflective of the time allocation patterns in other states. There may be unique characteristics of the policy environment of Tennessee that may not be the same in other states. For instance, Tennessee was a Race to the Top grant winner, which provided the resources necessary to build the TEAM evaluation system. Moreover, APs in Tennessee received training on the TEAM system that could have shaped the way APs use their time on instructional leadership.

II.8 Conclusion

This study provides some of the most systematic evidence to date about the factors that are associated with how APs allocate their time. It makes use of the rich data available in Tennessee to extend prior research on APs' time allocation. First, this study has data on a wide range of APs working across many different school contexts. Many of the previous studies have been limited to the members of professional organizations or urban APs (Austin & Brown, 1970; Glanz,

1994). This study is also the first study of APs' duties in the US to examine the relationship between school characteristics and the time use of APs. An especially important contribution of this study is that it examines how organizational characteristics like the number of support staff, principal characteristics, and time allocation of other leaders are related to how APs allocate their time.

This study suggests that how APs allocate their time has changed in the years since the last studies of APs' duties (Oleszewski et al., 2012; Sun, 2012). APs are not just dealing with buses, books, and balls. Today's APs are spending more time on instructional leadership and their time is intricately tied to the contexts in which they work. The key contribution of this study is in providing a sense of the magnitude of the differences in time allocation associated with various personal and school factors. For example, this study finds that the grade level of a school explains larger differences in time allocation than the evaluation ratings of principals. Additionally, the difference in time allocated to instructional duties associated with the gender of the AP are larger than the differences associated with the achievement level of the school. Female APs allocate about 4.5 percent more time to instructional duties than male APs, but an AP in a school one standard deviation above average in achievement only allocates 1.6 percent more time to instructional duties than an AP in a school of average achievement. These patterns suggest that variation in APs' time to duties may have a larger portion that is fixed based on a few stable characteristics of APs and their schools, and a smaller portion that is flexible based on the specific needs of a particular school.

More research is needed in several areas to extend this study. For example, research should examine how support staff can be allocated to schools to enhance or supplement the work of APs more effectively. Another fruitful area of research would be investigating the relationship

between how APs allocate their time and how prepared they are for the principalship. Does it matter if APs spend less time on administrative duties or discipline for their performance as principals? Additionally, research could explore whether there are equity concerns with how APs allocate their time. Since this study finds that gender and race of APs are related to how APs allocate their time, research should explore whether there are consequences to systematic differences in APs' time allocation. Perhaps the most important consideration for future research is whether the way APs allocate their time is related to how effective they are as APs and to student outcomes.

Chapter II Figures

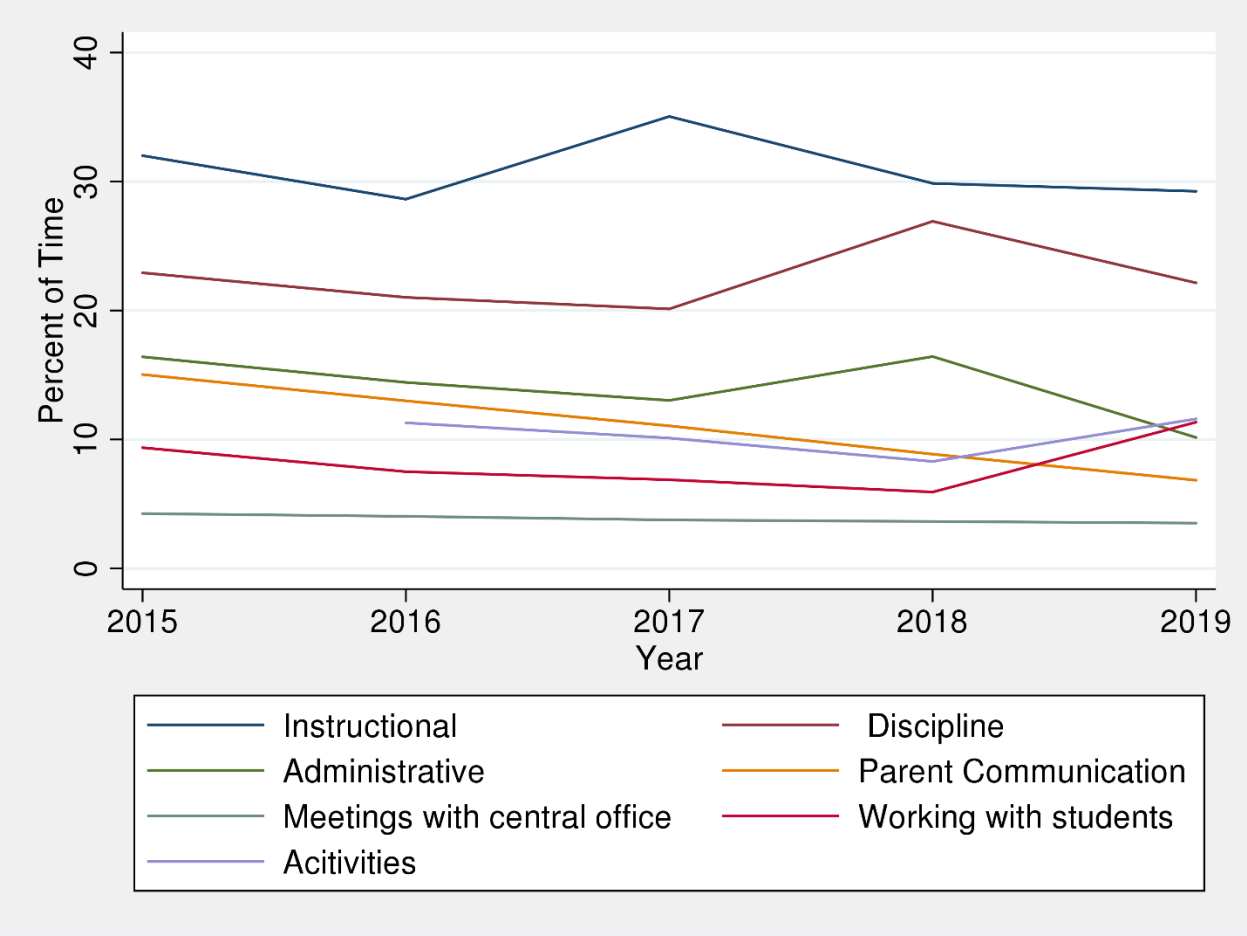


Figure II.1: Percent of APs' time spent on leadership tasks over time

Notes: The points represent the average amount of time APs reported spending on a given duty each year. There are 5756 AP-by-year observations included in this analysis.

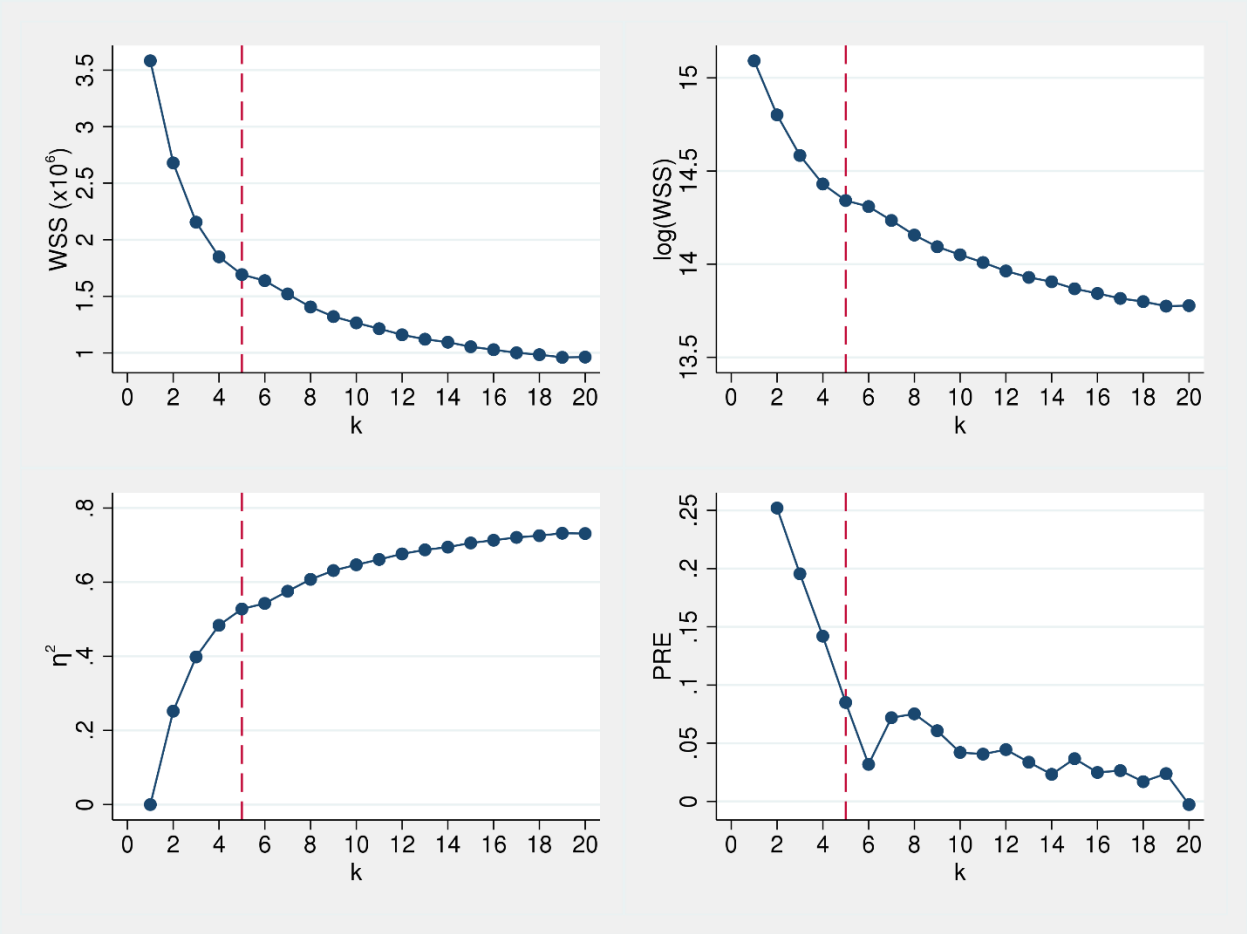


Figure II.2: WSS, log(WSS), η 2 , and PRE for all K cluster solutions

Notes: *k* refers to the number of clusters. This analysis is modeled after the analyses described in Makles (2012).

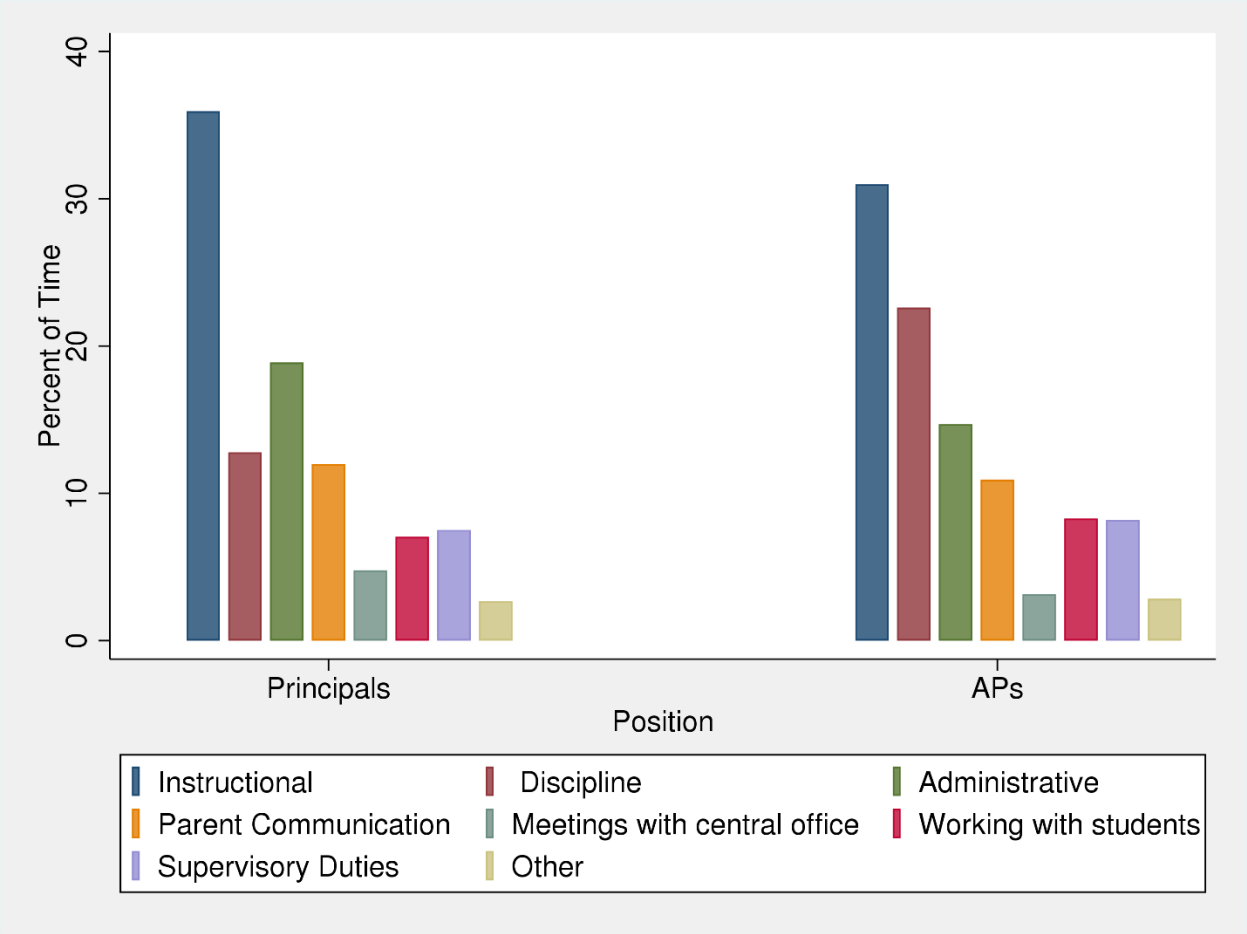


Figure II.3: Time use by position

Note: The bars represent the average percent of time allocated to a given duty among all person-by-year observations for APs and principals. There are 4,164 principals in the sample and 4,320 APs.

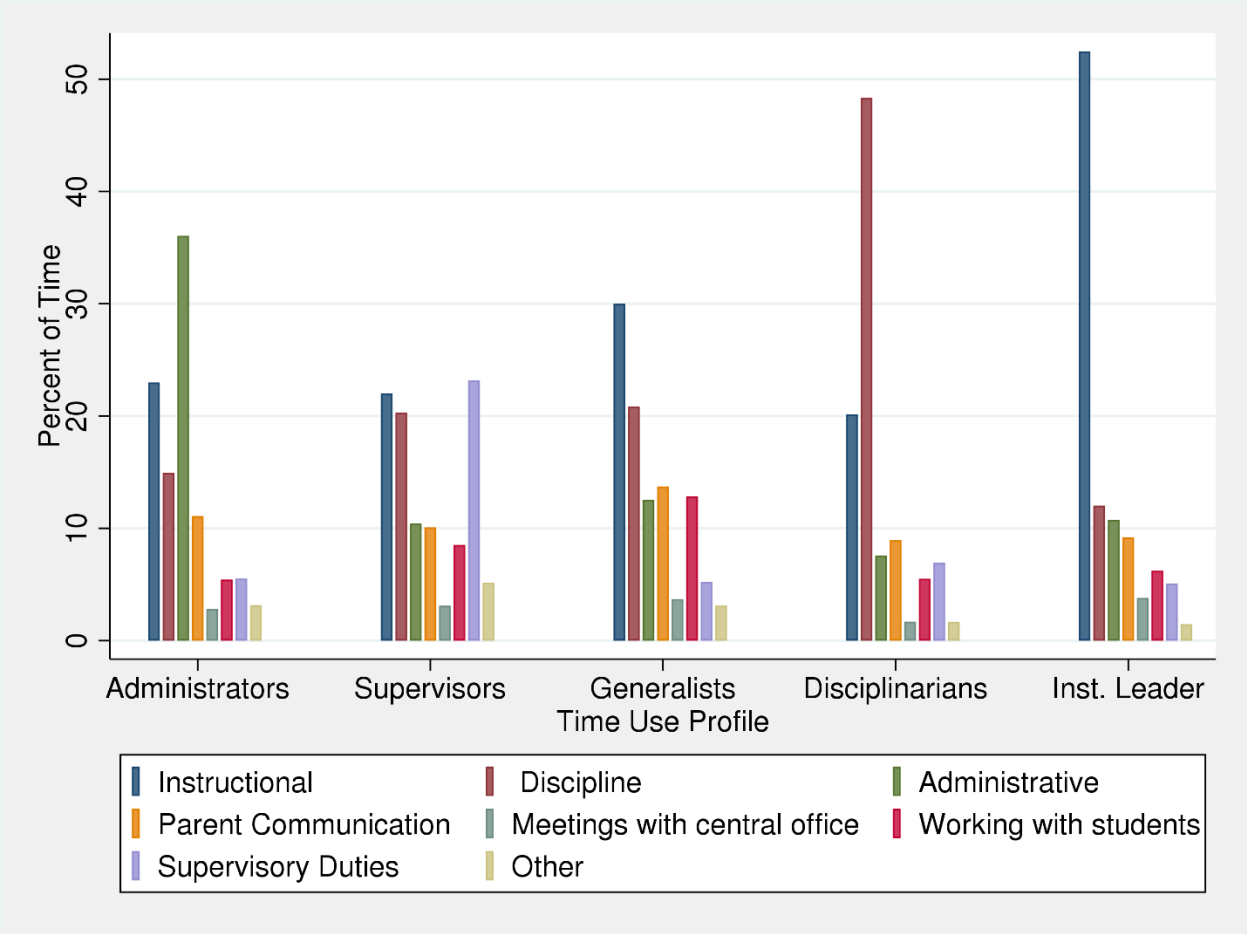


Figure II.4: Time use by AP time use profile

Note: The bars represent the average percent of time allocated to a given duty among all AP-by-year observations that are classified under a particular time use profile.

Chapter II Tables

Table II.1: Characteristics of sample APs as compared to all other APs

	Sample APs	Other APs
Female	0.61	0.53***
Black	0.13	0.24***
Age	46.08	46.50**
Years as an AP	4.01	2.89***
Years as a Teacher	13.49	15.94***
Highest Degree Master's	0.38	0.43***
Highest Degree Doctorate or Specialist	0.56	0.52***
Average Evaluation Rating	3.90	3.84***
N	4184	9644

Note: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.2: AP time use by clusters of time use

	Generalists	Inst. Leaders	Disciplinarians	Administrators	Supervisors
Discipline	20.83	12.01***	48.33***	14.95***	20.3
Instructional Leadership (index)	29.99	52.47***	20.15***	22.99***	22.01***
Administrative	12.53	10.75***	7.57***	36.04***	10.43***
Parent Communication	13.71	9.19***	8.93***	11.08***	10.09***
Meetings with Central Office	3.7	3.79	1.69***	2.82***	3.1**
Work with Students	12.84	6.22***	5.51***	5.43***	8.52***
Supervisory Duties	5.22	5.09	6.93***	5.52	23.17***
Other	1.18	0.48***	0.89	1.17	2.39***
N	1286	951	745	682	634

Notes: Clusters were created using the instructional leadership index rather than the individual tasks within the instructional leadership domain to maintain consistency across years. Percent of time allocated to other duties was only collected on the 2018 and 2019 TES.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.3: Characteristics of APs and their schools by time use cluster

	Generalists	Inst. Leaders	Disciplinarians	Administrators	Supervisors
Personal Characteristics					
Female	0.63	0.76***	0.51***	0.63	0.43***
Black	0.13	0.18**	0.12	0.08***	0.09**
Other race/ethnicity	0.01	0.02*	0.01	0.01	0.01
Age	45.65	46.55*	46.47*	45.4	46.56*
Master's Highest	0.37	0.39	0.4	0.37	0.4
EdD or EdS	0.58	0.55	0.54	0.57	0.56
Years as an AP	3.95	3.89	4.05	3.83	4.48**
Years as a Teacher	13.08	13.86**	13.75*	13.42	13.58
Evaluation Rating	3.92	3.92	3.79***	3.92	3.93
School Characteristics					
Proportion Black Students	0.19	0.24***	0.19	0.17*	0.15***
Proportion Latinx Students	0.1	0.11**	0.09	0.08*	0.07***
Proportion Other Race Students	0.04	0.04	0.04	0.05*	0.04
Proportion FRPL Eligible	0.54	0.56	0.58***	0.48***	0.52*
Proportion with IEPs	0.15	0.15	0.15	0.14**	0.14***
ADM	844.51	789.47**	878.19	899.57*	946.95***
Achievement Index	0.04	0.06	-0.01*	0.06	0.06
Elementary	0.37	0.53***	0.21***	0.31**	0.09***
Middle	0.22	0.15***	0.33***	0.17*	0.25
High	0.3	0.22***	0.39***	0.39***	0.52***
Urban	0.26	0.29	0.23	0.21*	0.17***
Suburban	0.21	0.21	0.16**	0.29***	0.18
Town	0.18	0.18	0.28***	0.16	0.19
Rural	0.35	0.32	0.33	0.33	0.45***
Principal's Evaluation Rating	3.91	3.93	3.89	3.95	3.93
Count of Support Staff	11.17	10.39**	12.01**	11.27	11.7
N	1148	822	697	605	603

Notes: N in this analysis is smaller than the N of time use because not all schools have achievement scores.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.4: Time use of APs by grade level

	Elementary	Middle	High
Time Use Profile			
Administrators	0.14	0.12	0.17*
Supervisors	0.04	0.17***	0.23***
Generalists	0.34	0.30*	0.26***
Disciplinarians	0.11	0.27***	0.20***
Inst. Leaders	0.36	0.15***	0.14***
Duty taking up the most time			
Discipline Most	0.19	0.46***	0.39***
Instructional Most	0.73	0.46***	0.43***
Administrative Most	0.11	0.12	0.19***
Parent Communication Most	0.04	0.06*	0.05
Meetings with CO Most	0.00	0.00	0.00
Working with Students most	0.03	0.06**	0.07***
Supervisory Duties Most	0.01	0.05***	0.10***
Other Most	0.01	0.01	0.01
N	1325	899	1401

Notes: Proportion of APs who report spending the most time on categories do not sum to 1 because APs could potentially spend equal amounts of time on multiple duties.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.5: Time use of APs by locale type

	Urban	Suburban	Town	Rural
Time Use Cluster				
Administrators	0.14	0.22***	0.13	0.15
Supervisors	0.10	0.13	0.14**	0.19***
Generalists	0.32	0.30	0.27*	0.30
Disciplinarians	0.17	0.13*	0.25***	0.16
Inst. Leaders	0.27	0.22*	0.20**	0.20***
Duty taking up the most time				
Discipline Most	0.29	0.28	0.40***	0.32
Instructional Most	0.61	0.56*	0.49***	0.55**
Administrative Most	0.13	0.20***	0.13	0.15
Parent Communication Most	0.04	0.07**	0.07**	0.05
Meetings with CO Most	0.00	0.00	0.00	0.00
Working with students Most	0.05	0.05	0.05	0.05
Supervisory Duties Most	0.03	0.06*	0.05*	0.07***
Other Most	0.01	0.01	0.00	0.01
N	1026	898	842	1514

Notes: Proportion of APs reporting spend the most time on categories do not sum to 1 because APs could potentially spend equal amounts of time on multiple duties.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.6: Predicting percent of time on duties from characteristics of APs and their schools

	Inst	Disc	Admin	Parent	Mtgs CO	Stu	Super
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Personal Characteristics							
Female	4.526***	-3.321***	1.025+	0.123	0.015	0.259	-2.398***
	(0.627)	(0.788)	(0.552)	(0.331)	(0.136)	(0.396)	(0.327)
Black	1.417	0.340	-3.558***	1.159*	0.131	1.085*	-0.161
	(1.247)	(1.180)	(0.662)	(0.520)	(0.209)	(0.506)	(0.378)
Other Race/Ethnicity	3.882+	-1.764	0.245	-0.647	0.389	-0.091	-1.073+
	(2.224)	(2.309)	(1.309)	(0.886)	(0.465)	(1.125)	(0.644)
Age	0.057	0.016	-0.107**	0.022	0.001	0.012	-0.015
	(0.050)	(0.062)	(0.037)	(0.025)	(0.013)	(0.026)	(0.023)
Years as an AP	-0.030	-0.038	-0.070	-0.063	-0.021	0.090	0.169**
	(0.114)	(0.121)	(0.100)	(0.086)	(0.030)	(0.062)	(0.060)
Years as a Teacher	-0.076	0.121+	0.016	-0.031	-0.028	-0.056+	0.064+
	(0.061)	(0.070)	(0.041)	(0.032)	(0.018)	(0.030)	(0.033)
Years in Same Position	-0.010	-0.030	0.164	-0.038	-0.009	0.018	-0.091
	(0.134)	(0.139)	(0.119)	(0.086)	(0.033)	(0.069)	(0.066)
Ed.D. or Ed.S. highest	-0.181	0.145	-0.173	-0.101	0.015	0.477+	-0.279
	(0.616)	(0.648)	(0.443)	(0.283)	(0.115)	(0.281)	(0.270)
School Characteristics							
Proportion Black	6.697**	-3.318	1.833	-1.870+	-0.345	-0.567	-2.500*
	(2.457)	(2.952)	(1.825)	(1.092)	(0.450)	(0.851)	(1.011)
Proportion Latinx	13.235**	-8.310+	1.602	-3.437**	1.470*	0.788	-4.889**
	(4.160)	(4.426)	(2.230)	(1.232)	(0.631)	(1.356)	(1.692)
Proportion Other Race	-2.453	4.031*	-2.520+	2.883**	-0.997*	-2.623	1.096
	(1.834)	(1.955)	(1.404)	(0.914)	(0.446)	(1.752)	(0.825)
Proportion FRPL	-3.728+	7.471**	-3.496+	-0.095	-0.444	-1.350	1.509
	(2.216)	(2.229)	(2.021)	(0.997)	(0.600)	(0.999)	(1.086)
Proportion IEPs	20.645*	19.726+	-21.273***	-1.100	-0.896	2.840	-17.899***
	(9.327)	(11.035)	(5.515)	(3.798)	(1.858)	(3.015)	(3.764)
Enrollment in 100s	-0.037	0.098	-0.008	0.040	-0.037+	-0.040	-0.014
	(0.084)	(0.079)	(0.062)	(0.040)	(0.021)	(0.039)	(0.054)
Achievement Index (Std)	1.640**	-1.292*	-0.839+	0.120	0.033	0.119	-0.039
	(0.557)	(0.609)	(0.502)	(0.258)	(0.118)	(0.318)	(0.354)
Middle School	-6.408***	8.296***	-2.823***	-0.244	-0.231	0.004	1.435***
	(0.867)	(1.075)	(0.466)	(0.395)	(0.184)	(0.471)	(0.364)
High School	-6.298***	3.615***	-0.261	-1.385**	0.191	1.131**	2.826***
	(1.113)	(0.988)	(0.586)	(0.420)	(0.199)	(0.413)	(0.616)
Suburban	-0.330	-0.969	1.184	0.712	-0.315	-0.131	-0.052
	(0.927)	(0.866)	(0.939)	(0.455)	(0.221)	(0.446)	(0.479)
Town	-0.896	2.623*	-1.136	0.131	-0.045	0.018	-0.362
	(1.145)	(1.254)	(0.832)	(0.522)	(0.262)	(0.450)	(0.528)
Rural	0.047	-0.589	-0.456	0.050	0.189	-0.023	0.838+
	(0.991)	(0.872)	(0.678)	(0.443)	(0.230)	(0.417)	(0.432)
Adjusted R-squared	0.16	0.10	0.05	0.16	0.17	0.05	0.37
Observations	3417	3417	3417	3417	3417	3417	3417

Notes: Standard errors clustered by district in parentheses. Column headers denote the outcomes for the models. All models include year fixed effects. Disc refers to discipline, Inst refers to instructional, Admin refers to administrative, Parent refers to parent communication, Mtgs CO refers to meetings with central office personnel, Stu refers to working directly with students, and Super refers to supervisory. The achievement index is standardized within year and is logically imputed for schools that do not have them in any given year using the linear midpoint between the two adjacent years.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.7: Predicting AP time use profile from characteristics of APs and their schools

	Generalists	Inst. Leader	Disciplinarian	Administrators	Supervisors
	(1)	(2)	(3)	(4)	(5)
Personal Characteristics					
Female	0.019	0.086***	-0.044**	0.023	-0.085***
	(0.015)	(0.014)	(0.016)	(0.016)	(0.015)
Black	0.028	0.031	0.016	-0.075***	-0.001
	(0.032)	(0.029)	(0.033)	(0.018)	(0.018)
Other Race/Ethnicity	-0.022	0.105+	-0.041	0.006	-0.048
	(0.069)	(0.063)	(0.068)	(0.051)	(0.056)
Age	0.000	0.003*	0.000	-0.002*	-0.001
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Years as an AP	-0.001	-0.001	0.002	-0.003	0.004
	(0.003)	(0.004)	(0.003)	(0.002)	(0.002)
Years as a Teacher	-0.003+	-0.001	0.002	0.001	0.002
	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)
Years in Same Position	-0.000	-0.001	-0.001	0.004	-0.002
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Ed.D. or Ed.S. highest	0.021	-0.004	-0.000	-0.001	-0.017
	(0.020)	(0.017)	(0.014)	(0.013)	(0.012)
School Characteristics					
Proportion Black	-0.064	0.187***	-0.080	0.038	-0.081*
	(0.052)	(0.055)	(0.061)	(0.051)	(0.040)
Proportion Latinx	0.056	0.420***	-0.214**	-0.089	-0.173**
	(0.087)	(0.114)	(0.079)	(0.057)	(0.059)
Proportion Other Race	-0.116*	-0.019	0.019	-0.015	0.130**
	(0.054)	(0.043)	(0.039)	(0.036)	(0.039)
Proportion FRPL	-0.011	-0.076	0.124**	-0.091+	0.054
	(0.058)	(0.060)	(0.043)	(0.049)	(0.036)
Proportion IEPs	0.376+	0.285	0.456*	-0.506**	-0.612***
	(0.220)	(0.225)	(0.207)	(0.185)	(0.144)
Enrollment in 100s	0.004	-0.004*	0.001	-0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Achievement Index (Std)	0.000	0.045**	-0.035*	-0.016	0.006
	(0.019)	(0.016)	(0.016)	(0.013)	(0.019)
Middle School	-0.016	-0.144***	0.164***	-0.044**	0.041*
	(0.023)	(0.019)	(0.027)	(0.016)	(0.016)
High School	-0.065**	-0.111***	0.076**	0.014	0.086***
	(0.023)	(0.024)	(0.023)	(0.020)	(0.022)
Suburban	-0.011	0.013	-0.024	0.032	-0.010
	(0.032)	(0.027)	(0.025)	(0.025)	(0.019)
Town	-0.011	0.009	0.046	-0.009	-0.035+
	(0.027)	(0.032)	(0.032)	(0.022)	(0.021)
Rural	-0.011	0.024	-0.036	-0.006	0.030+
	(0.023)	(0.024)	(0.025)	(0.018)	(0.016)
Adjusted R-squared	0.03	0.10	0.07	0.02	0.13
Observations	3417	3417	3417	3417	3417

Notes: Standard errors clustered by district in parentheses. Column headers denote the outcomes for the models. All models include year fixed effects. The achievement index is standardized within year and is logically imputed for schools that do not have them in any given year using the linear midpoint between the two adjacent years.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.8: Predicting percent of time on duties from support staff in the school

	Inst	Disc	Admin	Parent	Mtgs CO	Stu	Super
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Combined							
Support Staff	0.172*	-0.052	0.053	-0.033	-0.029	-0.054	-0.085*
	(0.085)	(0.130)	(0.064)	(0.044)	(0.020)	(0.034)	(0.043)
Adj. R-Sq	0.16	0.10	0.05	0.16	0.17	0.05	0.37
N	3417	3417	3417	3417	3417	3417	3417
Panel B: Specific Positions							
Counselors	-0.290	-0.261	0.420	-0.312	0.130+	0.178	0.199
	(0.496)	(0.436)	(0.395)	(0.197)	(0.078)	(0.182)	(0.239)
Instructional	0.170+	-0.001	0.074	-0.032	-0.039	-0.038	-0.125**
	(0.096)	(0.142)	(0.078)	(0.057)	(0.025)	(0.039)	(0.045)
Materials	0.927+	-0.864	-0.206	0.381	0.182	-0.098	0.003
	(0.503)	(0.632)	(0.419)	(0.276)	(0.169)	(0.306)	(0.335)
Administrative	-1.467	-1.080	1.618	-0.367	0.513	0.345	-0.522
	(1.141)	(1.242)	(1.074)	(0.502)	(0.313)	(0.694)	(0.630)
Student	0.278	-0.188	-0.083	0.017	-0.063	-0.307**	0.092
	(0.186)	(0.225)	(0.126)	(0.076)	(0.052)	(0.114)	(0.125)
Adj. R-Sq	0.16	0.10	0.05	0.16	0.17	0.06	0.37
N	3417	3417	3417	3417	3417	3417	3417

Notes: Standard errors clustered by district in parentheses. Column headers denote the outcomes for the models. All models include APs' personal characteristics, school characteristics, and year fixed effects. Disc refers to discipline, Inst refers to instructional, Admin refers to administrative, Parent refers to parent communication, Mtgs CO refers to meetings with central office personnel, Stu refers to working directly with students, and Super refers to supervisory.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.9: Predicting AP time use profile from support staff in the school

	Generalists	Inst. Leader	Disciplinarian	Administrators	Supervisors
	(1)	(2)	(3)	(4)	(5)
Panel A: Combined					
Support Staff	-0.000	0.004+	-0.000	-0.001	-0.003
	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
Adj. R-Sq	0.03	0.10	0.07	0.02	0.13
N	3417	3417	3417	3417	3417
Panel B: Specific Positions					
Counselors	-0.010	0.003	-0.007	0.011	0.002
	(0.012)	(0.015)	(0.011)	(0.012)	(0.009)
Instructional	0.003	0.002	0.000	-0.002	-0.004*
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)
Materials	0.010	0.014	-0.012	-0.011	-0.001
	(0.020)	(0.012)	(0.018)	(0.013)	(0.012)
Administrative	-0.022	-0.022	0.025	0.043	-0.025
	(0.039)	(0.036)	(0.035)	(0.033)	(0.026)
Student	-0.016*	0.013**	-0.001	0.001	0.002
	(0.007)	(0.005)	(0.006)	(0.004)	(0.007)
Adj. R-Sq	0.03	0.10	0.07	0.02	0.13
N	3417	3417	3417	3417	3417

Notes: Standard errors clustered by district in parentheses. Column headers denote the outcomes for the models. All models include APs' personal characteristics, school characteristics, and year fixed effects.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.10: Predicting percent of time on duties from principal characteristics

	Inst	Disc	Admin	Parent	Mtgs CO	Stu	Super
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Principal: Female	-0.125	-0.066	0.013	0.228	-0.099	-0.030	-0.005
	(0.377)	(0.332)	(0.260)	(0.167)	(0.066)	(0.187)	(0.168)
Principal: Black	0.365	0.267	0.134	-0.022	0.166	-0.302	-0.438
	(0.615)	(0.772)	(0.550)	(0.349)	(0.148)	(0.368)	(0.343)
Principal: Evaluation Rating (Std)	2.724+	-3.645***	2.473***	-0.079	-0.069	-0.039	-1.086**
	(1.561)	(0.964)	(0.694)	(0.473)	(0.172)	(0.714)	(0.389)
Principal: Years as a Principal	-0.067	0.045	0.059	-0.028	-0.041	-0.077	0.116*
	(0.091)	(0.129)	(0.090)	(0.043)	(0.025)	(0.051)	(0.050)
Principal: Years in the Same School	0.187	-0.099	-0.134	-0.019	0.057+	0.176**	-0.179**
	(0.140)	(0.148)	(0.095)	(0.062)	(0.034)	(0.055)	(0.063)
Adjusted R-squared	0.17	0.11	0.05	0.17	0.16	0.06	0.38
Observations	3143	3143	3143	3143	3143	3143	3143

Notes: Standard errors clustered by district in parentheses. Column headers denote the outcomes for the models. All models include APs' personal characteristics, school characteristics, number of support staff, and year fixed effects. Disc refers to discipline, Inst refers to instructional, Admin refers to administrative, Parent refers to parent communication, Mtgs CO refers to meetings with central office personnel, Stu refers to working directly with students, and Super refers to supervisory. Principal evaluation ratings are standardized within each year.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.11: Predicting percent of time use profile from principal characteristics

	Generalists	Inst.Leader	Disciplinarian	Administrators	Supervisors
	(1)	(2)	(3)	(4)	(5)
Principal: Female	-0.011	-0.002	0.002	0.007	0.004
	(0.011)	(0.010)	(0.008)	(0.009)	(0.007)
Principal: Black	0.002	0.007	0.015	-0.008	-0.015
	(0.021)	(0.015)	(0.015)	(0.016)	(0.014)
Principal: Evaluation Rating (Std)	-0.016	0.079*	-0.078**	0.061**	-0.047*
	(0.028)	(0.037)	(0.025)	(0.020)	(0.018)
Principal: Years as a Principal	-0.004	-0.000	0.002	-0.002	0.004
	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
Principal: Years in the Same School	0.004	0.005	-0.003	-0.002	-0.004
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)
Adjusted R-squared	0.03	0.10	0.07	0.02	0.13
Observations	3143	3143	3143	3143	3143

Notes: Standard errors clustered by district in parentheses. Column headers denote the outcomes for the models. All models include APs' personal characteristics, school characteristics, and year fixed effects. Principal evaluation ratings are standardized within each year.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.12: Predicting percent of time on major duties from other leaders' percent of time on duties

	Inst	Inst	Inst	Disc	Disc	Disc	Admin	Admin	Admin
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Other APs: Inst. Leadership Pct	-0.052		0.026	0.050		-0.081	0.036		0.058
	(0.046)		(0.058)	(0.053)		(0.057)	(0.032)		(0.039)
Other APs: Discipline Pct	-0.094		-0.110	0.025		0.032	0.030		0.054+
	(0.064)		(0.070)	(0.040)		(0.044)	(0.029)		(0.031)
Other APs: Administrative Pct	-0.027		0.000	0.050		0.039	-0.005		-0.005
	(0.048)		(0.066)	(0.038)		(0.057)	(0.046)		(0.051)
Principal: Inst. Leadership Pct		0.036	0.084+		0.100**	0.108*		-0.012	-0.072*
		(0.033)	(0.049)		(0.034)	(0.045)		(0.021)	(0.035)
Principal: Discipline Pct		0.175**	0.101		-0.026	-0.053		-0.053	0.026
		(0.053)	(0.095)		(0.047)	(0.088)		(0.037)	(0.082)
Principal: Administrative Pct		0.082*	0.112+		0.030	0.032		-0.038	-0.071
		(0.037)	(0.057)		(0.035)	(0.048)		(0.029)	(0.046)
Adj. R-sq	1316	1937	837	1316	1937	837	1316	1937	837
N	0.11	0.12	0.11	0.15	0.19	0.16	0.06	0.06	0.07

Notes: Standard errors clustered by district in parentheses. Column headers denote the outcomes for the models. All models include APs' personal characteristics, school characteristics, number of support staff, principal characteristics, district fixed effects, and year fixed effects. Disc refers to discipline, Inst refers to instructional, and Admin refers to administrative.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table II.13: Predicting percent of time on minor duties from other leaders' percent of time on duties

	Parent (1)	Parent (2)	Parent (3)	Mtgs CO (4)	Mtgs CO (5)	Mtgs CO (6)	Stu (7)	Stu (8)	Stu (9)	Super (10)	Super (11)	Super (12)
Other APs: Inst. Leadership pct	0.045*		-0.060	0.007		-0.079	0.004		-0.058	0.012		-0.126*
	(0.019)		(0.061)	(0.011)		(0.059)	(0.020)		(0.061)	(0.021)		(0.057)
Other APs: Discipline pct	0.044*		0.048	0.014+		0.045	0.049*		0.060	0.011		-0.005
	(0.021)		(0.045)	(0.008)		(0.049)	(0.022)		(0.044)	(0.020)		(0.045)
Other APs: Administrative pct	0.0336		0.054	0.006		0.046	-0.008		0.071	0.039		-0.001
	(0.027)		(0.063)	(0.009)		(0.058)	(0.025)		(0.056)	(0.026)		(0.054)
Other APs: Same as Outcome Duty pct	0.032		0.081	0.076		0.153	0.103+		0.084	0.195***		-0.121+
	(0.036)		(0.085)	(0.050)		(0.180)	(0.055)		(0.063)	(0.056)		(0.071)
Principal: Inst. Leadership pct		-0.013	0.122*		-0.001	0.124*		-0.024	0.125*		0.018	0.057
		(0.020)	(0.052)		(0.008)	(0.048)		(0.016)	(0.048)		(0.015)	(0.045)
Principal: Discipline pct		-0.006	-0.045		0.006	-0.028		-0.031	-0.044		0.026	-0.119
		(0.024)	(0.094)		(0.012)	(0.085)		(0.024)	(0.091)		(0.023)	(0.086)
Principal: Administrative pct		0.030	0.043		0.001	0.047		-0.005	0.049		0.001	-0.027
		(0.021)	(0.056)		(0.010)	(0.047)		(0.023)	(0.051)		(0.020)	(0.050)
Principal: Same as Outcome Duty pct		0.052+	0.040		0.069***	0.135		0.073*	0.092		0.191***	-0.161*
		(0.029)	(0.082)		(0.019)	(0.123)		(0.035)	(0.074)		(0.035)	(0.065)
Adj. R-sq	1316	1937	837	1316	1937	837	1316	1937	837	1316	1937	837
N	0.16	0.18	0.16	0.17	0.16	0.16	0.05	0.06	0.16	0.43	0.40	0.17

Notes: Standard errors clustered by district in parentheses. Column headers denote the outcomes for the models. All models include APs' personal characteristics, school characteristics, number of support staff, principal characteristics, district fixed effects, and year fixed effects. Parent refers to parent communication, Mtgs CO refers to meetings with central office personnel, Stu refers to working directly with students, and Super refers to supervisory.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Chapter II Appendix Tables

Table II.A1: AP counts and response rates in TES by year

	2015	2016	2017	2018	2019	Total
Counts	879	779	826	860	935	4279
Response Rate	47%	41%	43%	42%	49%	

Table II.A2: Duties surveyed by year of TES

	Duties	2015	2016	2017	2018	2019
Instructional	Instructional planning	Y	Y	Y	C1	
Instructional	Observing teachers	Y	Y	Y	C1	
Instructional	Coaching teachers	Y	Y	Y	C2	
Instructional	Providing observation feedback			Y	C2	
Instructional	Modelling lessons			Y		
Instructional	Teacher evaluations					Y
Instructional	Other instructional leadership					Y
	Administrative duties	Y	Y	Y	Y	Y
	Discipline	Y	Y	Y	Y	Y
	Working directly with students	Y	Y	Y	Y	
	Attending meetings with or sponsored by central office	Y	Y	Y		Y
	Monitoring extra-curricular activities		Y	Y	Y	
	Parent or community concerns		Y	Y	Y	Y
	Supervisory roles					Y
	Other student related meetings					Y
	Personnel matters					Y
	Other				Y	Y

Note: C1 and C2 indicate responsibilities that were combined in the 2018 TES

Table II.A3: List of common position titles for support staff

Position Title	Freq.	Percent
Adult Basic Education Personnel	163	0.09
Adult Education Personnel	9	0.01
Assessment Personnel	317	0.18
Attendance Staff	31	0.02
Attendance Teacher	94	0.05
Audiologist	99	0.06
Audiologist/Hearing Specialist	52	0.03
CTE Supervisor	21	0.01
Chapter 1 Supervisor	55	0.03
Community Services	46	0.03
Counselor	2,276	1.28
Family Resource Centers	41	0.02
Federal & Special Pgms	233	0.13
Federal Supervisor	8	0
Finance	6	0
Food Service Staff	6	0
Food Services	3	0
Home/Hospital Instruction	223	0.13
Homeschool Instructor	7	0
Human Resources	15	0.01
Instructional Coach	360	0.2
Instructional Materials/Technology Pe..	2,486	1.4
Interventionist (IRT)	510	0.29
Junior ROTC	1,258	0.71
Librarian	66	0.04
Librarian (Elementary/Secondary)	719	0.41
Librarian (Elementary)	7,070	3.98
Librarian (Secondary)	2,714	1.53
Maintenance	3	0
Materials Supervisor	10	0.01
Military Service	21	0.01
Military Service with CL	16	0.01
Non-Instructional and other Support S..	88	0.05
Other Instructional - System Wide	1,609	0.91
Other System Wide w/wo CL	2,378	1.34
Physical Education	2,137	1.2
Physical Education Teacher (Elementary)	22,052	12.43
Pre and After School Care	26	0.01
Reading Specialist	3,390	1.91
Sc Food Svc Supv	23	0.01
Sch Improvement & Accountability	100	0.06
School Counselor (Elementary)	13,841	7.8

Table II.A1 cont.

Position Title	Freq.	Percent
School Counselor (Elementary/Secondary)	1,470	0.83
School Counselor (Secondary)	11,400	6.42
School Curriculum Coord	2,925	1.65
School Curriculum Coordinator	111	0.06
School Health Coordinator	107	0.06
School Improvement & Accountability	5	0
School Nurse	15	0.01
School Psychologist	1,280	0.72
Social Worker	764	0.43
Special Ed School Psychologist	431	0.24
Special Education Options 7,8,9	8,037	4.53
Special Education Related Services	2,111	1.19
Special Education Supervisor	129	0.07
Special Education Teacher	6,473	3.65
Special Education Teacher (Elementary)	46,309	26.09
Special Education Teacher (Secondary)	24,842	14
Speech Specialist	77	0.04
Speech/Hearing Specialist	4,950	2.79
Student Data Management	8	0
Student Information Svcs	34	0.02
Superintendent	147	0.08
Supervisor Of Instruction (Elementary)	117	0.07
Supervisor Of Instruction (Elementary..	119	0.07
Supervisor Of Instruction (Secondary)	126	0.07
Supervisor of Instruction	31	0.02
Technology	185	0.1
Technology Staff	149	0.08
Testing Services	23	0.01
Testing and Assessment Personnel	28	0.02
Transportation	27	0.02
Vocational Supervisor	193	0.11
Total	177,205	100

Table II.A4: Reliability coefficients for items in the scale for instructional leadership

	2015	2016	2017	2018	2019
Instructional planning	0.33	0.40	0.53	†	
Observations			0.51	†	
Coaching	0.15	0.24	0.48		
Observe and give feedback	0.39	0.45	0.40		
Model lessons			0.57		
Teacher evaluations					††
Other instructional leadership					††
Overall	0.37	0.45	0.56	0.41	0.13

Note: Reliability coefficients for individual items cannot be estimated when there are only two items.
† indicates items included in the 2018 index of time spent on instructional leadership tasks
†† indicates items included in the 2019 index of time spent on instructional leadership tasks

Table II.A5: Correlation matrix of percent of time principals allocate to duties

	Inst	Disc	Admin	Parent	Mtgs CO	Stu	Super
Inst	1.000						
Disc	-0.315	1.000					
	(< 0.001)						
Admin	-0.406	-0.193	1.000				
	(< 0.001)	(< 0.001)					
Parent	-0.242	-0.007	-0.021	1.000			
	(< 0.001)	(0.004)	(< 0.001)				
Mtgs CO	-0.069	-0.138	-0.111	0.032	1.000		
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)			
Stu	-0.166	-0.029	-0.255	-0.106	-0.030	1.000	
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)		
Super	-0.299	-0.105	-0.137	-0.212	-0.096	-0.058	1.000
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	

Notes: Pair-wise correlations are reported in each cell with p-values in parentheses. Other duties was only measured in 2017-18 and 2018-19. Disc refers to discipline, Inst refers to instructional, Admin refers to administrative, Parent refers to parent communication, Mtgs CO refers to meetings with central office personnel, Stu refers to working directly with students, and Super refers to supervisory.

Table II.A6: Correlation matrix of percent of time other APs allocate to duties

	Inst	Disc	Admin	Parent	Mtgs CO	Stu	Super
Inst	1.000						
Disc	-0.452	1.000					
	(< 0.001)						
Admin	-0.210	-0.330	1.000				
	(< 0.001)	(< 0.001)					
Parent	-0.140	-0.120	0.004	1.000			
	(< 0.001)	(< 0.001)	(0.874)				
Mtgs CO	0.091	-0.247	-0.012	0.081	1.000		
	(< 0.001)	(< 0.001)	(0.614)	(0.001)			
Stu	-0.142	-0.167	-0.216	-0.060	-0.028	1.000	
	(< 0.001)	(< 0.001)	(< 0.001)	(0.013)	(0.251)		
Super	-0.252	-0.109	-0.179	-0.239	-0.122	-0.030	1.000
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(0.21)	

Notes: Pair-wise correlations are reported in each cell with p-values in parentheses. Other duties was only measured in 2017-18 and 2018-19. Disc refers to discipline, Inst refers to instructional, Admin refers to administrative, Parent refers to parent communication, Mtgs CO refers to meetings with central office personnel, Stu refers to working directly with students, and Super refers to supervisory.

Table II.A7: Correlation matrix of percent of time other APs and principals allocate to duties

		Principal						
		Inst	Disc	Admin	Parent	Mtgs CO	Stu	Super
Other APs	Inst	0.024	-0.059	-0.058	0.046	0.044	-0.001	-0.146
		(0.321)	(0.052)	(0.055)	(0.13)	(0.147)	(0.986)	(< .001)
	Disc	0.013	0.072	0.052	-0.047	-0.101	-0.038	-0.020
		(0.599)	(0.017)	(0.089)	(0.122)	(0.001)	(0.208)	(0.505)
	Admin	-0.030	0.044	0.038	0.149	0.013	-0.028	-0.048
		(0.223)	(0.15)	(0.218)	(< 0.001)	(0.682)	(0.352)	(0.118)
	Parent	-0.059	0.056	0.126	0.182	0.058	0.036	-0.204
		(0.016)	(0.068)	(< 0.001)	(< 0.001)	(0.055)	(0.242)	(< 0.001)
	Mtgs CO	0.002	-0.059	-0.062	0.104	0.299	0.035	-0.083
		(0.938)	(0.051)	(0.04)	(0.001)	(< .001)	(0.246)	(0.007)
	Stu	-0.030	-0.002	0.020	-0.037	0.050	0.091	-0.003
		(0.224)	(0.951)	(0.514)	(0.224)	(0.098)	(0.003)	(0.93)
	Super	0.035	-0.132	-0.105	-0.249	-0.058	-0.055	0.476
		(0.155)	(< 0.001)	(0.001)	(< 0.001)	(0.055)	(0.07)	(< 0.001)

Notes: Pair-wise correlations are reported in each cell with p-values in parentheses. Other duties was only measured in 2017-18 and 2018-19. Disc refers to discipline, Inst refers to instructional, Admin refers to administrative, Parent refers to parent communication, Mtgs CO refers to meetings with central office personnel, Stu refers to working directly with students, and Super refers to supervisory.

CHAPTER III

Preparing the Assistant Principal:

Preparation program features, perceptions, and effectiveness on the job

Research on the preparation of educators has expanded in recent years in the wake of efforts by government agencies to hold preparation programs accountable for educator quality (Cochran-Smith et al., 2015; Cochran-Smith & Villegas, 2015). Although much of the attention has been placed on teacher preparation programs (TPPs), school leadership preparation programs (SLPPs) have also received greater scrutiny in recent years (E. Anderson & Reynolds, 2015; Dorion & Reedy, 2018; Ni et al., 2017). As the evidence base regarding the importance of school leaders continues to grow, the urgency of preparing effective school leaders has become more apparent (E. Anderson et al., 2018; Liebowitz & Porter, 2019; Turnball et al., 2016). Although most states expect school leaders to receive formal training to become school leaders (E. Anderson & Reynolds, 2015), little evidence exists to suggest which programs and what aspects of those programs are most effective for building the capacities and knowledge school leaders need (Ni et al., 2017). One reason for the limited research on SLPP effectiveness is that there are several challenges unique to evaluating SLPPs that have yet to be resolved (Fuller & Hollingworth, 2018; Grissom, Mitani, et al., 2019; Ni et al., 2017). However, these challenges also point to the need for continued research on SLPPs to refine the methods and frameworks used in studying SLPPs.

One of the challenges in identifying the features of SLLPs that are most effective at preparing school leaders is that there is a lack of consensus on which outcomes should be used to

evaluate programs (Clifford, Larsen, Lemke, Chambers, & Swanlund, 2016a; Fuller & Hollingworth, 2014, 2018; Grissom et al., 2019). Most studies have used either proximate outcomes like graduates' perceptions as measured through surveys (Ni et al., 2019; Orr & Orphanos, 2011) or long-run outcomes such as student achievement when SLPP graduates become principals (Clifford et al., 2016b; Corcoran et al., 2012). The issue with using the proximal outcomes is that they may not reflect the actual performance of graduates on the job (Davis & Darling-Hammond, 2012). There are several theoretical and methodological issues with using long-run outcomes identified by prior research, including heavy data demands, challenges to isolating the effect of preparation on student achievement, and accounting for learning lost in the years after graduation (Clifford et al., 2016a; Fuller & Hollingworth, 2014, 2018; Grissom et al., 2019). Research is needed that can bridge the gap between the short-run perceptual outcomes and the long-run outcomes in studies of SLPPs' impacts.

Another important aspect of identifying the effective features of SLPPs is differentiating between the effects of selection of aspiring leaders into SLPPs and the programmatic features of SLLPs on graduates' outcomes (Clifford et al., 2016a; Grissom, Mitani, et al., 2019; Orr & Orphanos, 2011). The primary concern is that there is non-random sorting of aspiring school leaders to SLPPs, and their outcomes once they graduate may systematically differ because of sorting (Levine, 2005; Orr & Orphanos, 2011). To this point, several studies have provided empirical evidence that there is non-random sorting of aspiring leaders into SLPPs (Corcoran et al., 2012; Grissom, Mitani, et al., 2019). Most studies of SLPPs argue that selection is an essential component of SLPP quality (Cosner et al., 2015; Darling-Hammond et al., 2010; S. H. Davis et al., 2005; E. Wang et al., 2018), but differentiating between selection and other

educational features of SLLPs is critical to identifying how programs contribute to the preparation of leaders.

Another key limitation in the literature on SLPP effects is that SLPPs focus on the principalship and fail to emphasize the assistant principalship (APs) (Fuller & Hollingworth, 2018; Marshall & Hooley, 2006; L. Searby et al., 2017). Yet program graduates often have their first experiences as school leaders in AP roles (Allen & Weaver, 2014; Busch et al., 2012; B. W. Davis et al., 2017; Farley-Ripple et al., 2012; Fuller et al., 2018). Most school leaders spend multiple years as APs prior to becoming principals, and their experiences as APs can shape their leadership practices in ways that may not be attributable to SLPPs (Fuller & Hollingworth, 2018; Marshall & Hooley, 2006). In contrast to principals, who may lean more heavily on their experience as APs, early-career APs are more likely to lean on the preparation they received in SLPPs (Bastian & Henry, 2015; Clark et al., 2009; Fuller et al., 2018). Moreover, a recent study suggests that SLPPs can have a meaningful impact on APs' instructional leadership capacity (L. Searby et al., 2017). Since the research literature suggests that there are limitations to both the short-run and long-run outcomes of SLPPs for evaluating their effectiveness, examining the performance of SLPP graduates as APs may serve as an important intermediary outcome.

The limitations of the available research on SLPPs points to the need for more research that describes the quality features of SLPPs, separates the influence of selection into SLPPs and their core programmatic features, and provides evidence that links SLPPs' quality features to intermediary outcomes. Using rich data available in Tennessee, this study will address each of these key issues in the literature. I will specifically address the following research questions: First, to what extent do the quality features of SLPPs vary across programs preparing APs in Tennessee? Second, to what extent are quality features of SLPPs related to APs' perceptions of

the quality of their preparation? Third, to what extent are quality SLPP features related to APs' effectiveness ratings? Lastly, to what extent are the relationship between quality SLPP features and APs' effectiveness ratings sensitive to the inclusion of prior performance measures? This study will extend prior research on SLPPs by being the first study of which I am aware to explore the relationship between effectiveness ratings of APs and the features of SLPPs.

In the next section of this chapter, I provide an overview of the relevant literature on educator preparation. This review will include relevant studies that link teacher preparation to teacher outcomes and research on SLPPs' quality features using case studies, perceptions of graduates, and long-run outcomes of graduates as principals. Then, I provide a description of the data sources, measures, and methods to be employed in this study. The methods sections will provide an explanation for the use of a novel measure of SLPP quality features in this study. Next, I describe the findings of this study and provide a discussion of the findings. Lastly, I discuss the limitations to this study and what future research should be conducted to address these limitations.

III.1 Literature Review

Research on SLPPs has grown in recent years as empirical evidence has mounted that school leaders play a pivotal role in student learning (E. Anderson & Reynolds, 2015; Young et al., 2009; Young, 2015a). Policymakers and researchers have increased their attention on SLPPs as they wrestle with how to increase the supply of effective school leaders (E. Anderson et al., 2018; E. Anderson & Reynolds, 2015). This attention on SLPPs is reflected in recent briefs written by professional organizations like the University Council on Educational Administration (UCEA), philanthropic organizations like the Wallace Foundation, and the US Department of

Education (S. H. Davis et al., 2005; US Department of Education, 2016; E. Wang et al., 2018; Young et al., 2012). Part of this increased research emphasis on SLPPs is in response to recent research that has been critical of SLPPs, particularly traditional university based preparation programs, and their ability to effectively prepare school leaders for the complexities of school leadership (Levine, 2005; Murphy & Vriesenga, 2006). These studies find that school leaders feel underprepared for their roles and are frustrated with the emphasis on theoretical rather than practical aspects of school leadership. The research community has responded to these critiques by concentrating research efforts on measuring and improving the quality of SLPPs (E. Anderson et al., 2018; Cosner, 2019; Young et al., 2012; Young, 2015a). Despite this emphasis on SLPPs in education leadership and policy research, rigorous quantitative studies on the preparation of school leaders has been relatively limited (Clifford et al., 2016a; Fuller & Hollingworth, 2018; Grissom, Mitani, et al., 2019; Orr, 2010; Pounder, 2012).

A few empirical studies have shaped the field's understanding of SLPPs and especially what makes an effective program (S. H. Davis et al., 2005; Ni et al., 2017; Orr & Orphanos, 2011; Young & Rorrer, 2012). This research has been informed by the research evaluating teacher preparation programs (TPP) and the research on what principals do to improve student learning (Darling-Hammond et al., 2010; Gates et al., 2014; Liebowitz & Porter, 2019; E. Wang et al., 2018). However, few studies examine how SLPPs prepare graduates for their roles as APs, even though most SLPP graduates become APs before becoming principals. In this section I will provide a review of the most relevant research on the features of TPPs that are related to teacher effectiveness, research on effective SLPPs, specific features of SLPPs identified in the literature as important to the preparation of aspiring leaders, and what is known about APs and their preparation.

III.1.1 Effective preparation of aspiring teachers

Although research on TPPs is not directly applicable to SLPPs, studies evaluating TPPs provide some important insights and considerations when thinking about how to evaluate SLPPS. There is a well-established base of quantitative research related to the evaluation of TPPs, including evaluations of specific programs and features of programs (Borko, 2004; Cochran-Smith, 2004; Hoffman et al., 2015). Cochran-Smith and Colleagues review of the recent research on TPPs featured in the *Handbook of Research in Teaching* (2016) and in the *Journal of Teacher Education* (Cochran-Smith et al., 2015; Cochran-Smith & Villegas, 2015) provide a broad and detailed synthesis of the literature in this area. Within the research on TPPs, I focus my literature review on quantitative analyses that relate TPPs programmatic features to graduates' outcomes. According to Cochran-Smith et al. (2016), much of this research has centered on examining how alternative preparation routes (Darling-Hammond et al., 2005; Xu et al., 2011) and assessment criteria (Goldhaber et al., 2017) have improved the quality of teachers entering the profession. A few studies have looked specifically at all the preparation programs within a specific labor market and the features of those programs (D. J. Boyd et al., 2009; Henry et al., 2014; Ronfeldt et al., 2013; Ronfeldt & Campbell, 2016). These studies of specific markets for TPPs have been identified as especially fruitful for the evaluation of TPPs because they allow for comparison of program features within a common policy environment (Cochran-Smith et al., 2016).

Evaluation of TPPs and their features reveal that there are small but measurable differences between programs, and differences in the preparation of pre-service teachers are significantly related to differences in their outcomes (Hoffman et al., 2015; Noell et al., 2019). Among the programmatic features of TPPs examined in quantitative studies, the design of the

curriculum and field experiences of TPPs have been found to be especially important to the preparation of teachers (Cochran-Smith et al., 2016). Aspects of the curriculum in TPPs that are significantly related to graduates' outcomes include the depth of the content-specific training, coherence across courses, and the integration of content with field experiences (D. J. Boyd et al., 2009; Grossman et al., 2008; Struyven et al., 2010). Field experiences have also received substantial research emphasis, and these studies find that aspiring teachers are best prepared through methods like field experiences that emphasize situated learning (Hoffman et al., 2015). Moreover, field experiences have been found to be especially impactful when mentor teachers are more effective and have access to well-established structures of support (Hoffman et al., 2015; Ronfeldt, 2012; Ronfeldt et al., 2018).

The literature on SLPPs has similar strands of research as the research on TPPs. Multiple studies of SLPPs have examined alternative preparation routes for preparing aspiring school leaders (Clifford et al., 2016b; Corcoran et al., 2012; Gates et al., 2014). Researchers have compared the programs within a single state to determine if there are measurable differences between programs (Fuller & Hollingworth, 2016; Grissom, Mitani, et al., 2019). A few studies have examined how specific features of SLPPs are related to the outcomes of program graduates (Ni et al., 2019; Orr & Orphanos, 2011). The research evaluating TPPs also identifies issues that may be relevant to evaluations of SLPPs. Research on TPPs suggests that studies of SLPPs should be especially mindful of the following: how state regulations constrain variation in the features of programs (D. Boyd et al., 2008), differences in graduates' outcomes may be relatively small (D. J. Boyd et al., 2009; Ronfeldt & Campbell, 2016), and field placements may be especially important to preparation (Hoffman et al., 2015; Ronfeldt et al., 2018).

III.1.2 Overview of issues and approaches to identifying effective SLPPs

Concomitant with the growth of research evaluating TPPs, the number of evaluations of traditional university based SLPPs started to grow. Among these new studies evaluating SLPPs, several critiques of traditional university based SLPPs have been raised (E. B. Goldring & Sims, 2005; Hess & Kelly, 2005, 2007; Levine, 2005; Murphy & Vriesenga, 2006). These criticisms include the lack of innovation, empirical research, connection to local districts, and relevant content in SLPPs. Levine (2005) finds that 89 percent of the principals surveyed in his study believe that schools of education do not adequately prepare them to cope with the realities of their profession. Moreover, Levine (2005) finds that the curriculum of most programs are fairly similar to each other, covering similar content areas. Hess and Kelly's (2007) analysis of the course syllabi from a sample of SLPPs across the country finds that school leaders receive limited exposure to content related to utilizing data, managing human capital, and dealing with parents. Murphy and Vriesenga (2006) in their review of the literature on the preparation of school leaders find that there is a dearth of empirical evidence on effective practices and features of SLPPs.

Recent studies have varied approaches to identifying quality SLPPs. The first type of research on quality SLPPs are case studies of SLPPs. These types of studies include the School Leadership Study (Darling-Hammond et al., 2010; S. H. Davis et al., 2005) and examinations of UCEA's Exemplary Preparation Programs (Cosner, 2019; Young et al., 2012; Young, 2015a). A second set of studies have examined outcomes of SLPP graduates to identify characteristics and features of programs associated with better graduate outcomes (Clark et al., 2009; Fuller et al., 2016; Ni et al., 2019; Orr & Orphanos, 2011; Pounder, 2012). A third set of studies have evaluated specific alternative programs (Clark et al., 2009; Corcoran et al., 2012; Gates et al.,

2014) or compared a set of programs within a specific market for SLPPs (Clifford et al., 2016b; Fuller & Hollingworth, 2016; Grissom, Mitani, et al., 2019) using outcomes of graduates. Each of these categories of research on SLPPs provide their own contributions and have limitations in examining how SLPPs prepare aspiring leaders for their roles.

III.1.3 Case Studies of Exemplary SLPPs

One of the most widely cited studies related to the preparation of school leaders is the School Leadership Study (Darling-Hammond et al., 2010; S. H. Davis et al., 2005). In this study, the researchers conduct eight case-studies of programs that are identified as exemplary and innovative by the SLPP literature, a panel of experts, and a survey of professional associations. The authors study these programs and their graduates to examine how these exemplary programs prepare school leaders and what evidence exists to suggest that they shape the knowledge and practices of their graduates. Davis and colleagues (2005) find that the content, methods, and structure of SLPPs are especially important to the quality of SLPPs. In terms of content, the School Leadership Study finds that exemplary programs have research-based and coherent curricula. The methods that the exemplary programs share are the implementation of problem-based learning, cohort groups, mentors, and field-based internships. Lastly, exemplary programs have structures in place for collaboration with local district partners.

In the years following the release of the findings from the Stanford School Leadership Study, UCEA initiated several programs aimed at increasing and improving the research on SLPPs (E. Anderson et al., 2018; Young et al., 2012; Young, 2015a). UCEA established the Exemplary Educational Leadership Preparation (EELP) Award and the *Journal of Research on Leadership Education* (JRLE). It also developed and implemented the School Leadership

Preparation and Practice Survey Instruments, which would later become the INSPIRE survey suite (Ni et al., 2019; Pounder, 2012; Winn et al., 2016). UCEA also published two editions of the *Handbook of Research on the Education of School Leaders* (Young et al., 2009; Young & Crow, 2017). These efforts resulted in a dramatic increase in the amount of research directed towards identifying effective programs and uncovering the practices of SLPPs that are central to the preparation of school leaders (E. Anderson et al., 2018; Ni et al., 2017; Young & Rorrer, 2012).

Among the studies that emerged out of the JRLE were several case studies of exemplary SLPPs, namely case studies of the EELP award winners. Analyses of the practices of EELP award recipients suggest that exemplary programs share many key qualities and practices (Cosner, 2019). EELP award recipients tend to be purposeful in selection of aspiring leaders into their programs (Cosner et al., 2015; Fusarelli et al., 2019). EELP programs are designed around providing aspiring leaders with opportunities to engage in powerful learning experiences (Cunningham et al., 2019; Young, 2015b). These powerful learning experiences include field-based internships that are intentional about placing aspiring leaders from EELP programs in contexts that expose them to diverse students and opportunities to engage in leadership practices (Cosner et al., 2012, 2015; Fusarelli et al., 2019). The partnerships that EELP programs have with local districts are built on shared values and incorporate feedback from district partners in all elements of the program design (Korach et al., 2019). Although the emphases of EELP programs may differ, they are all marked by a rigorous approach towards critical self-examination and continuous improvement (Cosner, 2019; Honig & Donaldson Walsh, 2019).

There are many similarities across these case studies of "exemplary" SLPPs. Some of the common features across programs include: rigorous selection processes, structures to ensure

quality interactions with peers such as cohort models, relevant and coherent curriculum, intensive and extensive field experiences, and collaboration with district partners in the design of all SLPP programmatic features. Although effective SLPPs share many similar features, it is unclear whether the similarities between SLPPs is a function of a consensus on the quality features of SLPPs or isomorphism. Anderson et al. (2015) suggests that state policies governing SLPPs demonstrate substantial isomorphism and may limit the variability in programs across the spectrum of quality. Levine (2005) also finds that there is little variation in the programmatic features of SLPPs and largely concludes that these similarities are part of the problem with university based programs. It is unclear whether this sort of isomorphism is a positive sign of programs responding to research, or programs reifying features that do not have proven connections to graduate effectiveness.

III.1.4 Quality features of SLPPs as identified by graduates' outcomes

Relatively little quantitative research exists that links SLPPs' programmatic features to the outcomes of SLPP graduates. Fuller and colleagues in their 2011 and 2016 studies of SLPP graduates and long-run outcomes like school achievement and labor market outcomes find that the Carnegie classification of the university that a principal graduates from is related to positive outcomes. Specifically, these studies find evidence to suggest that graduates of research-intensive universities tend to lead schools with higher achievement and have better job placement. Clark and colleagues' (2009) study of principals find some suggestive evidence that graduating from a more selective university is associated with higher student achievement. A problem with the Carnegie classifications and Barron's selectivity measures is that they are coarse and are not direct measures of program quality (Fuller et al., 2011, 2016).

Studies that have finer-grained measures of SLPPs' quality features use survey data from graduates and program directors to estimate measures of latent quality. Orr and Orphanos (2011), using structural equation modeling (SEM), find that SLPPs identified in the Stanford Principal Study as an exemplary program are related to both higher levels of leadership learning and better instructional leadership practices as a principal. Moreover, they find that programs with more quality features and better internships are related positively to leadership learning and leadership practice. Ni and colleagues (2019) examine programmatic features in their study using SEM to measure the latent quality of features and find that important program features identified in the literature (rigor and relevance of the program, internship, peer relationships, cohort structure, and faculty quality) are all positively and significantly related to graduates' knowledge of leadership. Of these aspects, the rigor of the program and faculty quality have especially large relationships with leadership knowledge, and faculty quality's relationship to learning is mediated by the program's rigor and relevance. These studies provide evidence of the specific features of SLPPs that may be especially important for graduate effectiveness in leadership roles, but these studies are largely limited because they do not directly measure the effectiveness of graduates in leadership roles and rely on survey data that can be cumbersome to collect. Orr and Orphanos (2011) do have a measure of leadership practices as measured by teacher surveys, but teacher evaluations of school leaders are only one of many measures used by states and districts to evaluate school leaders. Moreover, they purposefully sampled the programs identified as exemplary in the Stanford study, and these exemplary programs are more likely to have the quality features in their programs and more effective graduates.

III.1.5 Effective SLPPs as identified by graduates' outcomes

Evaluations of SLPPs face different theoretical and methodological challenges from evaluations of TPPs when using graduates' outcomes (Clifford et al., 2016a; Fuller & Hollingworth, 2014, 2018; Grissom, Mitani, et al., 2019). The first of these issues with using graduates' outcomes to evaluate SLPPs is that there is an ongoing debate about which outcomes should be used to evaluate SLPPs (Grissom, Mitani, et al., 2019). Outcomes examined in recent studies include short-run outcomes like graduates' survey responses (Barakat et al., 2019; S. H. Davis et al., 2005; Ni et al., 2019; Orr & Orphanos, 2011) and long-run outcomes like labor market outcomes, effectiveness ratings, and student achievement (Clifford et al., 2016b; Corcoran et al., 2012; Fuller & Hollingworth, 2018; Gates et al., 2014; Grissom, Mitani, et al., 2019). Second, SLPP graduates often do not start their careers as school leaders until several years after graduation, which creates the potential for learning to be lost or replaced in the intervening years (Clark et al., 2009; Fuller & Hollingworth, 2018). Lastly, it is unclear what covariates should be included when estimating the impact of SLPPs because there is evidence that aspiring leaders both systematically sort into SLPPs and systematically sort into schools when they graduate from SLPPs (Fuller et al., 2019; Fuller & Hollingworth, 2016; Grissom, Bartanen, et al., 2019; Grissom, Mitani, et al., 2019). Prior studies have largely thought of selection into programs as a key feature of SLPPs, but researchers have suggested that not adjusting for covariates in estimating the effect of programs on graduates may lead to perverse incentives that encourage programs to discourage their graduates from serving in low-performing schools (Fuller & Hollingworth, 2016). Despite these challenges, recent quantitative studies evaluating SLPPs have substantially advanced the field's understanding of the relationship between SLPPs and their graduates' outcomes (E. Anderson et al., 2018; J. Davis, 2016; Ni et al., 2017; Young, 2015a).

Some of these studies find no difference or mixed results between preparation and graduates' outcomes, while others find small but measurable differences. In a study of the relationship between different characteristics of principals and their impact on student achievement, Clark and colleagues (2009) find that there was no measurable relationship between the selectivity of the graduate program a principal attended and mixed results for New York City's training program for aspiring principals. Corcoran and colleagues (2012) find no relationship between attending the same New York City training program and student achievement using a slightly different sample. Clifford and colleagues (2016b) in their evaluation of SLPPs in the Alliance to Reform Education Leadership find that there are no measurable differences in student achievement of schools led by principals who were trained in an Alliance SLPP as compared to schools led by principals in other SLPPs. In their examination of the growth trajectories of schools and the characteristics of principals, Bowers and White (2014) find mixed results for the selectivity of the graduate program a principal attended. They find that selectivity of a principals' graduate programs is related to student achievement growth trajectories in schools outside of Chicago, but not for schools in Chicago.

A few studies do find measurable differences in graduates' outcomes among SLPPs, but these studies point to many potential issues related to evaluating programs using graduates' outcomes. Gates et al. (2014) find that students in schools led by principals prepared by the New Leaders Program have higher levels of student achievement as compared to similar schools with principals not prepared by the New Leaders Program. Although there are differences in student achievement for schools led by New Leaders principals, this study only finds measurable differences under certain conditions. For example, the relationship in elementary schools is found for students who have attended a school led by a New Leaders principal for two or three

years, and in high school there is only a difference for students in schools led by a New Leaders principal that has been at the same school for three years. Fuller et al. (2016) find that there are measurable differences in the placement rates of SLPP graduates, based on the production rates of SLPPs and the Carnegie classification of the university. However, their analysis suggests that the demographic characteristics of graduates may explain more variation in placement rates into leadership positions than their program characteristics. Grissom, Mitani, et al. (2019) find that there are measurable differences between program graduates across a number of outcomes that may be potentially useful in evaluating preparation programs, including placement rates, student achievement, and the survey responses of teachers. A concern raised by this study is that the relative rankings of programs are sensitive to the outcome used and the inclusion of covariates in the estimation procedure.

Perhaps the most challenging issue that many of these studies acknowledge is that there is a substantial amount of time that lapses between when aspiring leaders graduate from SLPPs and when they become principals. In these intervening years, many of the graduates become teacher leaders and APs, where they may gain knowledge and skills that may supplant what graduates learned in their SLPPs. Studies that examine shorter-run outcomes and survey responses of SLPP graduates have been able to look more closely at specific features of SLPPs, but these studies do not provide much evidence that quality features of SLPPs are related to actual changes in behavior and performance (S. H. Davis et al., 2005; Levine, 2005; Ni et al., 2019; Orr & Orphanos, 2011).

III.1.6 Preparing APs

Little is known about what it takes to prepare APs for their specific roles as school leaders (Allen & Weaver, 2014; Busch et al., 2012; Gurley et al., 2015; Marshall & Hooley, 2006). Most studies of APs' preparation suggest that their formal preparation is insufficient and does not adequately prepare them for the transition from teacher to school leader (Barnett et al., 2012; Oliver, 2005; Parylo et al., 2013). Searby and colleagues (2017) find evidence to suggest that APs who receive training directly aimed at preparing them to be instructional leaders feel more ready to be instructional leaders. Allen and Weaver (2014) find that APs who are recent graduates from SLPPs in Kentucky express the need for additional supports in almost every area of school leadership after starting their roles. The study finds that some of the areas that APs feel the least prepared for are in maximizing time spent on quality instruction, gathering and analyzing data about the educational environment, and assessing emerging trends in order to adapt leadership strategies (Allen & Weaver, 2014). Busch and colleagues (2012) interviewed recent SLPP graduates in Texas who were serving as APs and asked them what advice they would give aspiring school leaders in SLPPs. The authors suggest this data is an indicator of the areas that SLPPs can improve to better prepare APs. The areas that APs mention the most in their advice is gaining knowledge of curriculum and instruction, developing skills in fostering caring relationships with teachers, and working on being people who engender trust.

The research literature consistently finds that most school leaders have their first experiences in leadership as APs (Folsom et al., 2015; Fuller et al., 2016; Gates et al., 2004; Ringel et al., 2004; Turnbull et al., 2016). Although the literature consistently emphasizes this reality, the research on SLPPs and the learning of APs remains relatively thin (Marshall & Hooley, 2006; Oleszewski et al., 2012). Research on the roles and duties of APs suggests that

they are most likely to be asked to manage school discipline and to evaluate the instruction of teachers (Hausman et al., 2002; Sun, 2012), but studies of SLPPs suggest that the training APs receive may not adequately prepare them to be instructional leaders (Allen & Weaver, 2014; Barnett et al., 2012; Busch et al., 2012) and they receive almost no training to specifically serve as APs (Marshall & Hooley, 2006; Oleszewski et al., 2012). Without intentional emphasis on training APs through their experiences in SLPPs, the field will continue to ignore an important and large group of school leaders. The existing research does not explore the extent to which the skills needed to be an effective AP differ from the skills needed to be an effective principal. The research also provides little guidance as to what leads to the preparation of APs who are more likely to enter the principalship.

III.1.7 Gaps in the Literature and Contributions

This study attempts to bridge the literature on SLPPs by finding the connection between the quality features of SLPPs and measures of effectiveness as early-career school leaders. The performance of APs is an important intermediary outcome that can link the learning of SLPP graduates to their performance as principals. Prior case studies of SLPPs are helpful in generating hypotheses about the quality features of SLPPs that are most likely to be impactful to the effective training of aspiring school leaders, but there are only a few studies that find evidence to suggest that these quality features are related to graduates' outcomes (S. H. Davis et al., 2005; E. Wang et al., 2018). Prior quantitative studies of quality SLPP features have linked the features of SLPPs to the short-run outcome of graduates' leadership learning, but they have not connected these features to external measures of performance that are used for educator accountability (Ni et al., 2019; Orr & Orphanos, 2011). Studies linking specific SLPP programs

to long-run performance outcomes like student achievement and effectiveness as principals have taken the important step of demonstrating a link between SLPPs and important outcomes (Clifford et al., 2016b; Corcoran et al., 2012; Fuller et al., 2016; Gates et al., 2014; Grissom, Mitani, et al., 2019). However, these studies have encountered many methodological challenges, and have not been able to link student achievement or other long-run outcomes to SLPPs' quality features. Moreover, they do not differentiate between selection and other programmatic features of SLPPs or account for potential learning loss or replacement, and none of them have explicitly studied how SLPPs prepare their graduates to be APs.

This study is the first to link SLPP features to both the perceptions of graduates and their effectiveness ratings as APs. Since the research on SLPP features shows clear links to graduates' learning (Ni et al., 2019), and the research on the impact of specific SLPPs finds links to graduates' performance as school leaders (Gates et al., 2014; Grissom, Mitani, et al., 2019), this study bridges these two strands of quantitative research on SLPPs. Although we cannot directly link AP performance and student learning, evidence suggests that there is a link between experience as an AP and performance as a principal (Bastian & Henry, 2015; Clark et al., 2009). In addressing this primary gap in the SLPP research literature, this study will also provide analysis of a novel method for measuring the quality features of SLPPs using data that is not collected through surveys. The final contribution of this study is that it explores how the relationship between SLPPs and graduates' outcomes are explained by selection into programs.

III.2 Tennessee Policy Context

Tennessee is an ideal site to study SLPPs impacts on APs because the Tennessee State Board of Education (TSBE) had a re-accreditation process for all of Tennessee's SLPPs in 2014. This re-

accreditation process produced detailed information on the programmatic design of all of Tennessee's SLPPs. Additionally, TDOE's data has information on the SLPPs Tennessee's educators attended between 2011-12 and 2016-17, and this information can be linked to administrative personnel and evaluation files. Although there are some requirements in TSBE's educator preparation policy on common features that need to be included in each program, there is still substantial variation in programs. During the period of this study, there were 19 accredited in-state SLPPs in Tennessee, and all of the school leaders in this study's sample attended one of the 19 programs. Although in theory an aspiring school leader could be prepared through an alternative preparation program like New Leaders and get licensed through the out-of-state SLPP process, the state board has not accredited any out-of-state or alternative programs to prepare Tennessee's leaders. These programs produce different numbers of graduates each year, and the counts by year from the data are presented in Table III.1.

Two different Tennessee policies are especially relevant to this study, TSBE's educator preparation (5.504) and educator licensure (5.502) policies. These policies were adopted in 2014 and included updated standards for school leaders, requirements for school leader licensure, and expectations for educator preparation program accreditation. In response to the adoption of these policies governing educator preparation and licensure, the Tennessee State Board of Education required SLPPs to submit documents and artifacts to be reaccredited. The accreditation of TPPs in Tennessee are conducted primarily through the Council for the Accreditation of Educator Preparation (CAEP), and many of the expectations for SLPP accreditation mirror the CAEP expectations. All the universities with accredited SLPPs in Tennessee have CAEP accreditation for their TPPs. SLPP accreditation occurs on a five-year cycle, and the most recent accreditation was completed at the beginning of the 2014-15 school year. Although final re-accreditation was

approved in the 2014-15 school year, SLPPs submitted documentation and artifacts starting in the summer of 2014, and many of the submitted documents were created or revised in 2012 and 2013.

The program re-accreditation proposals were required by TSBE's educator preparation policy to contain evidence of how SLPP's are meeting expectations in the areas of content and pedagogical knowledge, clinical partnerships, candidate selection, and quality assurance. One of the changes that occurred as part of the 2014-15 re-accreditation process was the expectation that SLPPs would demonstrate that they were preparing their graduates to meet the Tennessee Instructional Leadership Standards (TILS). The TSBE prepared SLPPs of this increased emphasis on the TILS when evaluating SLPPs for approval. All of the SLPPs responded in slightly different ways to this change in standards, but most programs made changes to their programs to bring them into clearer alignment with the TILS.

At the end of the 2014-2015 accreditation process, there were 19 state certified SLPPs in Tennessee. These programs are distributed widely throughout the state, and some of the largest programs have satellite locations across the state. A map of the primary campuses of all 19 SLPPs is provided in Appendix Figure III.A1. This study focuses on the period after the 2014 re-accreditation because prior research suggests that the design of SLPPs is relatively stable unless policy requires SLPPs to change (E. Anderson & Reynolds, 2015; Levine, 2005). A reasonable assumption made in this study is that the design of SLPPs will be fixed between re-accreditation years. Therefore, program quality as measured in 2015 should be applicable to all program graduates between 2015 and 2019, the next accreditation year. Most program graduates are teachers at the time of completing their programs, and most graduates become APs for their first leadership role before becoming principals. Among the sample of SLPP grads from 2015 to

2017, 70 percent were still teachers, 20 percent were employed as APs, and 3 percent had some experience as principals by 2018. Figure 2 shows the proportion of SLPP grads in teaching, AP, and principal positions during the period of this study[15]. On average, APs in Tennessee have approximately 8.8 years as teachers before they graduate from SLPPs, and continue to work as teachers for a little less than five years before becoming APs. However, the median number of years of teaching experience SLPP grads have when graduating is seven years, and the median number of years SLPP grads continue to teach before their first AP job is three years. Figure 3 plots the average years since graduating from an SLPP for new APs and new principals.

III.3 Data and Measures

The data on Tennessee's SLPPs programmatic features come from two sources. The first source of data is the re-accreditation proposals SLPPs in Tennessee submitted to TSBE as described in the previous section. These data are paired with data that is publicly available through SLPPs' websites, which were coded between the fall of 2015 and the spring of 2016. The program websites provide some additional information on SLPPs not covered in the SLLP re-accreditation documents, especially in the area of selection. Although the literature on SLPPs quality features suggests that faculty quality is an important component of program quality (Ni et al., 2019; Young et al., 2012), information on the faculty of SLPPs from program websites and re-accreditation documents is limited and not consistently available across all SLPPs.

Information from these two data sources was coded using a system of binary codes that were analyzed using a partial independence item response (PIIR) approach that is described in further detail in the methods section. The coding and subsequent analyses of the codes produced

[15] Proportions do not sum to 1 because some SLPP grads leave TN public schools, work in other school-based positions, or work in central office in the years after graduation.

factor scores for SLPP quality along the dimensions of selection, curriculum, field experiences, and partnerships. The dimensions of SLPP quality measured in this analysis are a product of TSBE's policy on educator preparation that requires SLPPs to report on these specific areas. Although these components are not exhaustive of all the features that differentiate SLPPs on their quality, these are some of the core areas as described in the literature. To ensure the reliability of the coding, the documents have been independently coded by a second researcher and from them I estimate the average inter-rater reliability (Cohen's Kappa) and average agreement. There are no glaring discrepancies, therefore no reconciliation between coding is needed. Appendix Table III.A4 reports the inter-rater reliability estimates for our independent coding.

The next data source for this study comes from a survey of early-career APs who participated in the Tennessee Educator Survey (TES). The TES is an annual survey administered in partnership between the Tennessee Education Research Alliance and TDOE. The TES is offered to all public-school educators in Tennessee and asks participants to answer questions about their work, schools, and leadership. Each year the TES asks teachers and school leaders a separate set of core questions and assigns additional branch questions either at random or based on the characteristics of the educator as specified by the survey. The TES in 2018 and 2019 asked early-career school leaders (in their first three years as administrators) what SLPP they attended, what year they graduated from the SLPP, and their perceptions of the quality of their SLPP. These questions were answered by more APs than principals because most school leaders have their first experience in leadership as APs. Approximately 65 percent of participants who answered these questions were APs, and this fraction represents 27 percent of all eligible APs[16]

[16] The TES had a response rate of 58% in 2018 and 2019 for administrators. This response rate can be found on the 2019 TES report retrieved from https://www.tn.gov/content/dam/tn/education/data/2019-survey/Survey_Report.pdf.

(those that have three or fewer years as a school leader). The data from the TES are used to answer research question 2. The sub-sample of APs who responded to the TES provides a validity check for the measures of SLPP quality. If the SLPP quality measure is valid, APs trained at SLPPs with more quality features should have more positive perceptions of their preparation.

The final source of data that this study draws upon is the administrative data from TDOE obtained through the Tennessee Education Research Alliance at Vanderbilt University. These administrative data contain rich information on school personnel and students. The student data include information about the gender, race, FRPL, and IEP statuses. The personnel data contain information on individual's gender, race, age, and evaluation scores.

The evaluation system for Tennessee's school personnel is called the Tennessee Educator Acceleration Model (TEAM). Under the TEAM system, districts in Tennessee must evaluate all teachers and school leaders, but they can choose the specific framework and rubric they want to use, as long as the framework covers all of the state's standards. Both the overall system of evaluation and a specific rubric developed by TDOE are called TEAM, but for the purposes of this study when I refer to TEAM score I am referring to the broader evaluation system. TEAM evaluation scores for school leaders and teachers are made up of subjective practice ratings by a supervisor and measures of student achievement. Since prior research has raised concerns about using student achievement scores to evaluate SLPPs and school leaders (Fuller & Hollingworth, 2018; Grissom, Kalogrides, et al., 2015; Grissom, Mitani, et al., 2019), this study will focus on the practice ratings by supervisors. The ratings by supervisors can range from one to five, with indicators aligned to the Tennessee Instructional Leadership Standards. Indicator scores are then averaged to produce an overall average. Grissom, Blissett, and Mitani (2018) find in their study

of the TEAM supervisor ratings for school leaders that principals' scores across the domains of the TEAM rubric measure one underlying factor of leadership effectiveness. They also find that the factor scores are relatively stable over time and predictive of student achievement and other measures of principal quality. Ronfeldt and Campbell (2016) find that TEAM ratings of teachers could be useful for evaluating the program design of teacher preparation programs and have a significant relationship with value-added scores. Based on this prior research, the ratings of SLPPs' graduates should be an appropriate outcome to evaluate SLPPs' features. The evaluation policy in Tennessee requires APs to be evaluated multiple times, so I use the average ratings of APs within a year.

Table III.2 compares the sample of new APs who are included in this study to all other new APs. Table III.3 describes the schools they worked in the first year they worked as APs. The sample of APs for this study graduated between the years 2015 and 2017, are within three years of graduation, and are in their first two years as APs. Table III.4 compares SLPP graduates across cohorts and these results suggest that there are few demographic differences between cohorts of SLPP graduates. The only observable difference between cohorts is degree attainment, which is likely due to the shorter time frame for SLPPs to earn terminal degrees for later cohorts. I only include APs who graduated from their SLPP between 2015 and 2017 because the documents describing SLPPs quality features are from the accreditation process in 2014. The sample only includes APs who are within three years of their graduation because there is the potential that what graduates learned in their SLPPs may degrade over time. I restrict the sample to APs in their first two years as APs because their on-the-job experience may supplant what they learned in their SLPP after the first two years. I include the evaluation rating from the first two years of their careers as APs rather than just their first-year evaluation ratings to account for

the potential for some measurement error in a single year of evaluation scores. By averaging the evaluation ratings of APs in their first two years when available, I am able to reduce some of the noise that may be associated with a single year of evaluation ratings. However, it may be that the preparation an AP received from an SLPP is most relevant to their first year as a school leader, so I replicate the analyses using only the first-year evaluation ratings. To answer research question four, I also use the average evaluation ratings of SLPP's graduates as teachers. I average the evaluation ratings of SLPP's graduates as teachers over all available years prior to an individual's year of graduation from an SLPP.

One area of tension that exists in both the TPP and SLPP literature is how to define a program. Some studies have defined each individual degree or licensure program offered by a university as separate while others have grouped all programs within a university together (E. Anderson et al., 2018; Ronfeldt & Campbell, 2016). The decision to define a program in one or another way has implications for the estimation of the quality features measures. If degree and licensure programs within the same institution have substantive differences in quality features, grouping the programs together will reduce the variation in programs. If programs in the same institution are relatively similar, separating out programs may artificially weight more heavily the importance of features found in institutions with multiple programs. Since SLPPs in Tennessee are accredited at the university level and the accreditation documents do not distinguish between degree and licensure programs, I define a program as a university. If there are substantive differences across programs within the same university, I am unable to observe those differences based on the document analysis and my measure of quality features may be noisier. Fortunately, few programs have a licensure-only program, and most of the program graduates earn a master's degree rather than a doctorate or specialist degree.

III.4 Methods

III.4.1 Coding of program features

The measures of the quality features are produced from a coding of the SLPP re-accreditation documents and program websites. Program websites were coded between the fall of 2015 and the spring of 2016. The coding scheme is largely based on prior research on the quality features of SLPPs (Darling-Hammond et al., 2010; Ni et al., 2019; Orr & Orphanos, 2011; Young, 2015a), but I also allow for codes to emerge inductively from the data. Each SLLP is coded for quality features in the areas of selection, curriculum, field experiences, and partnerships. The codes for selection quality include minimum requirements for admissions, materials required with applications, and any interviews or approvals required. The codes for curriculum quality include codes for the presence of relevant leadership courses, the number of relevant courses, and assessments of knowledge (Grissom & Loeb, 2011; Sebastian et al., 2018). The codes for field experiences include the length, the SLPP's involvement in coordinating the internship, the process for choosing mentors, and the products expected from the internship. Lastly, the codes for the district partnerships are based on the extent to which partner districts are involved in each of the other three domains of SLPP quality features. For example, a SLPP will have a higher quality feature score for district partnerships if partners are involved in the selection process.

The method employed for estimating a program's quality scores is modelled after the procedure that Strunk (2011) employs for coding district collective bargaining agreements (see also Strunk et al., 2018; Strunk & Grissom, 2010). Strunk (2011) uses a generalized form of Reardon and Raudenbush's (2006) Partial Independence Item Response (PIIR) method to code collective bargaining agreements using a series of binary items. These binary codes indicate whether a provision that restricts the behavior of school districts is included in the contract. The

PIIR model allows for the possibility that the presence of one provision is partially dependent on the presence of another provision. In the case of the coding of SLPP accreditation documents, if an SLPP requires recommendation letters for their selection process, then the code for recommendation letters is coded as a one. If that same SLPP requires two recommendation letters, the code for a second recommendation letter is coded as a one as well. The coding for the final reduced set of items used in estimating the PIIR scores are reported in Table III.8 and Appendix Tables III.A1-III.A3. The reduced set of items included in the estimation is determined by excluding items in the measure that had reliability coefficients of less than .25.

The PIIR model is based on Item Response Theory (IRT), which estimates the true ability scores of test takers based on their test item responses. The IRT procedure predicts a latent ability score for test takers based on which answers the test taker got correct, giving more weight to correct answers on questions that most test takers missed. In the PIIR context of this study, each SLPP receives a latent quality features score for each of the four SLPP features of selection, curriculum, internships, and district partnerships. This estimation procedure predicts the quality score based on how many codes for high quality practices are coded as one through the website and document review. This PIIR procedure gives more weight to the presence of practices that most SLPPs do not have. This latent quality score can be thought of as the underlying quality of an SLPP's features, with higher scores meaning more high-quality features are included in that domain of the SLPP.

In this study I specifically employ PIIR to develop a measure of the quality features of SLPPs because prior studies have faced challenges to identifying variation in SLPP quality. Prior studies that have compared SLPPs have been unable to identify what specifically makes some SLPPs more effective than others (Clifford et al., 2016b; Corcoran et al., 2012; Gates et al.,

2014; Grissom, Mitani, et al., 2019). Other studies have used proxies for SLPP quality such as Carnegie classifications or Barron's selectivity (Clark et al., 2009; Fuller et al., 2016). The problem with these coarse measures is that they are likely to contain substantial noise and do not address the reality that most school leaders are not trained at research intensive universities, limiting the relevance of such a measure (Hackmann & McCarthy, 2013).

A couple quantitative studies have evaluated a large number of features of SLPPs and have used procedures to account for the relative noise to signal in their measures (Ni et al., 2019; Orr & Orphanos, 2011). These studies are notable in that they use structural equation modeling to account for the complex relationships between different features of SLPPs by creating latent variables and estimating their relationship to graduates' outcomes. Although these studies are more sophisticated, they rely on survey data that are gathered from program directors, and this survey collection is onerous and difficult to collect. As more states and districts integrate survey tools like the INSPIRE suite to regularly evaluate SLPPs, the challenges associated with collecting this data may be less of an issue. For the time being it may be relatively difficult to collect the rich data from the INSPIRE surveys. Since the PIIR method utilized in this study only relies on documents that are collected as a part of the established accreditation process and from program websites, there is no additional cost on the part of states or programs to collect the data. These procedures should provide a replicable, statistically based procedure for measuring the variation across SLPPs. Moreover, this procedure does not rely on subjective decisions for the inclusion of specific features in defining the quality of programs, and instead uses empirically based processes for deciding to include specific features in the final estimation of quality features. The result should be a reliable and valid measure that differentiates SLPPs based on their programmatic features.

III.4.2 Descriptive Analyses

The analyses for this study will begin with a descriptive examination of the characteristics of the SLPPs in Tennessee. This analysis describes the variation in quality features across the SLPPs and the specific features that contribute the most to the overall latent quality scores. The latent quality scores estimated using the PIIR approach are standardized to be relative to a theoretical distribution of quality based on the sample of SLPPs in Tennessee. In addition to a descriptive analysis of the quality features of SLPPs, I provide analyses of the characteristics of SLPP graduates (gender, race, age, years of experience) and the schools that they work in as school leaders.

III.4.3 Regression Analyses

To answer the second and third research questions, I conduct a straightforward OLS regression that models perceptions of program quality or evaluation ratings of APs as a function of SLPP latent quality controlling for a number of demographic characteristics of APs and their schools. This model is written formally in Equation 1:

$$y_{it} = \beta_0 + \sum_1^k \beta_k SLPPQual_{ikt} + X'_{it}\Gamma' + \tau_t + \epsilon_{it} \quad (1)$$

y_{it} represent the outcome measure (perceptions or evaluation rating) of AP *i* from graduating cohort *t*'s. $SLPPQual_{ikt}$ is a vector of the PIIR quality measures of *k* SLPP features experienced by an AP. X_{it} represents a vector of covariates included in these regressions as controls that may be related both to the measured latent quality and APs' perceptions. These controls include the AP's gender, race, educational attainment, and years of experience as a teacher. School covariates include school size, grade level, and locale type of the school (urban, suburban, town,

or rural). I also add controls for the proportion of students in the school that are Black, Latinx, FRPL eligible, and have IEPs.

To answer the final research question, I assess whether any of the significant β_k in Equation 1 are sensitive to attempts to account for the effect of non-random sorting of individuals to SLLPs. The strategy that I employ to answer this question is adding controls that proxy for the pre-SLLP leadership capacity of SLPP graduates. I use the average effectiveness ratings of SLLP graduates who were teachers in years prior to individuals graduating from their SLLP as a proxy for pre-SLPP effectiveness. The use of this measure as a proxy for the pre-SLLP leadership capacity is based on a theory which suggests that effective teachers have stronger instructional knowledge that can be tapped into as school leaders (Hallinger & Murphy, 1985), and recent empirical evidence which suggests that effectiveness as a teacher may be related to effectiveness as a principal (Goldhaber et al., 2019). If the inclusion of these covariates causes the relationship between the quality of SLPP features and AP effectiveness ratings to disappear, then non-random sorting of individuals to SLLPs would explain most of the differences in graduates' effectiveness. The standard error estimates for these regression analyses are robust to heteroskedasticity and clustering at the level of SLPPs. I cluster standard errors at the SLPP level because assignment to different levels of SLPP quality is a function of the SLPP that an AP attended (Abadie et al., 2017).

III.5 Findings

III.5.1 Descriptive Analyses

The findings from the descriptive analyses in this section address the first research question of this chapter. Table III.5 presents the characteristics of SLPP graduates from each program that

became an AP within three years after graduation. Although we are looking at a smaller sample of SLPP graduates, the overarching findings from this table are consistent with the findings in Grissom et al. (2019). In Grissom et al. (2019), the authors find that aspiring leaders tend to sort to SLPPs along observable demographic characteristics. The graduates of most programs are proportionally more female than male, but a few SLPPs have lower proportions of female graduates than the average. Notably Programs C, E, and N had closer to even numbers of male and female graduates. Aspiring leaders also sort to programs along racial lines with concentrations of Black SLPP graduates in a few programs. The majority of the graduates of Programs R and S are Black while programs A, G, H, and M did not produce any Black graduates during the study period.

Table III.6 suggests that SLPP graduates tend to sort into schools that they work in post-graduation. For example, the graduates of programs R and S tend to work in schools with majority Black student populations after graduation. The graduates of Programs F, G, and N tend to work in very large schools after graduation. Table III.6 suggests that this sorting of graduates to schools may in part be a function of the locales of their schools. For example, most of Program M's graduates work in rural schools and tend to work in smaller schools with less racially diverse student populations. Some of the most notable observations come from the counts of SLPP graduates included in the sample. A few programs produce a large number of the SLPP graduates who go on to become APs within three years of graduation. Program D did not produce any graduates who went on to become APs within three years of graduation. This low yield may in part be a reflection of the smaller number of graduates that Program D produces every year, and that the program tends to specifically cater to aspiring leaders of independent and religious schools.

Table III.7 reports the PIIR quality scores for Tennessee's SLPPs. Table III.7 presents the PIIR quality scores for each of the four coded domains and the overall score. The results in Table III.7 suggest that most of the PIIR program quality scores are related to one another, but the four domains do not overlap perfectly. The only domain score that does not seem to fit the general pattern of the other three domains is the field experiences domain. Appendix Table A III.3 has the coding for the items included in the estimation of the field experience PIIR scores and there is little variation in the items across programs. Most of the quality features are either present in almost all of the SLPPs or very few of the SLPPs. This lack of variation across items for the field experiences domain may suggest that the PIIR score for that domain may be less reliable. As a robustness check I run all of the analyses excluding the field experience quality scores and the results are generally similar. When looking across domains in SLPP quality scores, it stands out that three programs (A, G, and R) have consistently higher PIIR scores, and three programs (F, C, and B) are consistently at the bottom of PIIR scores. Program A has the highest score in selection criteria and quality of district partnerships, and it has the second highest score in rigorous curriculum. Program F has the lowest score on curriculum and partnerships and the second lowest score on selection. Program A and F have the same score for field experiences.

Table III.8 presents the coding for the selection domain of quality SLPP features. The SLPPS are sorted based on their Selection PIIR scores in ascending order. The results of the coding as reported in Table III.8 suggest that there is substantial variation in the type and number of selection criteria that SLPPs require for admissions. What seems to distinguish SLPPs with higher selection PIIR scores from SLPPs with lower scores is especially the academic criteria for SLPPs. SLPPs with more rigorous selection criteria tend to have higher expectations for prior academic achievement. Programs A and G have the highest selection scores, and these two

programs require minimum undergraduate GPAs and require a relatively high minimum score on either the GRE or the MAT. The results of the coding from the other three domains of SLPP quality are presented in Appendix Tables III.A1-A3.

These tables, when taken together, suggest that there are some consistent patterns across programs and some of the programs with the most quality features in one domain tend to have more quality features in other domains. What also stands out is that although Programs A, G, and R have the highest quality scores, they tend to only prepare a few SLPP graduates. The three programs only prepared 23 APs from 2015-2017. This points to the possibility that the smaller size of the SLPPs is part of what makes it possible for them to have more quality features. The three SLPPs with the lowest quality scores (F, B, and C) prepared many APs. Programs F, B, and C prepared a total of 112 APs during the same time period, and Program B prepared the most APs of all programs in Tennessee.

III.5.2 Regression Analyses

Table III.9 answers the second research question of this study. This table presents the regression results from models of graduates' perceptions of their SLPP's quality as a function of the SLPP's PIIR scores. The results in Table III.9 are the estimates when including controls for the APs' personal characteristics and school characteristics, but the coefficients on the SLPP's quality are fairly consistent whether the model controls for the characteristics of the APs, schools, or unobserved district heterogeneity. The results for models that include different covariates are included in Appendix Table A III.5. Since an AP's perceptions of their SLPP may be influenced by their working conditions at the time of taking the survey, I believe that the models that include controls for school characteristics are the most accurate estimates.

The results in Table III.9 suggest that the overall PIIR score does not have a measurable relationship with measures of graduates' perceptions of their SLPP's quality. However, the domain scores for curriculum, partnerships, and field experiences, are all significantly related to graduates' perceptions of their SLPP's quality. New school leaders who attended SLPPs with higher curriculum PIIR scores are more likely to say that their SLPP was rigorous and that they would attend their SLPP again. A one standard deviation increase in the curriculum PIIR score is associated with a .23 point increase in a new school leader's agreement with the statement that their program was rigorous and a .27 increase in their agreement that they would take the program again if they were given the choice. Similarly, SLPP graduates who attended a program with a one standard deviation higher score on field experiences were associated with a .26 points higher level of agreement that their SLPP was rigorous. Interestingly, new APs who attended SLPPs that had stronger district partnerships, holding all other characteristics constant, were less likely to report that they would attend their SLPP again and less likely to describe their SLPP as rigorous. SLPP graduates who attended a school one deviation higher than average in district partnerships would have a .47 lower agreement that their SLPP was rigorous and a .34 lower agreement that they would attend their program again. These results suggest that each of the PIIR scores of SLPP quality is a distinct domain with distinct relationships with graduates' perceptions of quality. It also suggests that the differences in graduates' perceptions associated with differences in SLPP PIIR scores is substantial. A standard deviation in the variable's scale is .63 and .68 for graduates' agreement that their program was rigorous and that they would take their program again, respectively. The estimates range from about one third to two thirds of a standard deviation.

Table III.10 presents the results of regressions of APs' evaluation ratings on SLPPs' quality scores. This table includes the estimates from models with no controls except for fixed effects for the graduation cohort, a model with controls for APs' personal characteristics, and a model with controls for school characteristics. The results from these models answer the third research question about the relationship between SLPP's features and early career performance as an AP. The outcome measure for these analyses is the average evaluation rating of the graduate as an AP from their first two years as an AP in the scale of the evaluation rating system. Scores in the evaluation system can range from one-five and the average AP evaluation rating has a mean of 3.73 points and a standard deviation of .55 points.[17] Table III.10 suggests that the overall PIIR quality score has no measurable relationship with APs' evaluation ratings. Among the specific domain scores, the quality scores for curriculum, partnerships, and field experiences are all significantly related to evaluation ratings of APs in their first two years. These are the same domains that are significant predictors of APs' perceptions of their SLPPs quality. However, the relationships between the PIIR scores and evaluation ratings are in the opposite direction of their estimated relationships with perceptions of SLPP quality. Based on the results from the fully saturated model in column 6, a one standard deviation increase in an SLPP's PIIR score in curriculum is associated with a .11 point decrease in graduates' average evaluation ratings as APs. A one standard deviation increase in an SLPP's field experience PIIR score is associated with a .28 point decrease in graduates' average evaluation ratings as APs. However, a one standard deviation increase in an SLPP's district partnerships score is associated with a .29 point increase in graduates' evaluation ratings as APs.

[17] As a sensitivity analysis I conducted all of these analyses with only the graduates' evaluation ratings in their first year as APs and all the results are substantively similar. These results are reported in Appendix Tables III.A6 and III.A7.

Another pattern that emerges from Table III.10 is that the estimates on the PIIR scores seem to change when covariates are included in the model. The estimate of selection PIIR scores in column 4 is -.02 when no covariates are included but -.07 in column 6 when both personal and school covariates are included in the model. The coefficient on curriculum PIIR scores becomes less negative when covariates are included in the model. While the substantive interpretation of the coefficients does not change across models, these changes in the coefficient related to the inclusion of covariates suggest that the relationship between the PIIR scores and evaluation ratings are somewhat sensitive to the inclusion of covariates.

Table III.11 addresses the fourth research question. The results from Table III.11 present the results from regressions estimating the relationship between APs' average evaluation ratings in the first two years and PIIR scores for SLPPs controlling for APs' prior performance as teachers. The measure of prior performance as a teacher is an average of teacher evaluation ratings in all years prior to graduating from their SLPP. This measure is in the original scale of the evaluation rubric, which ranges from one to five. The mean of the average evaluation ratings as a teacher is 3.95 points and has a standard deviation of .51 points. The first notable finding from this table is that graduates' evaluation ratings as APs are related to their evaluation ratings as teachers. This point estimate would suggest that the average teacher evaluation ratings of APs are a useful proxy for capacity to be effective principals, as measured by their AP evaluation ratings. The results are consistent with prior findings that suggest that performance as a school leader is related to measures of performance as a teacher. The estimates of the relationship between average rating as a teacher and average rating as an AP ranges from about .30 to .46 points depending on the model and covariates included. A one-point change in teacher evaluation ratings is equivalent to about two standard deviations, so a one-point change represents a

relatively large difference in evaluation ratings as a teacher. Interestingly, the adjusted R^2 suggests that the inclusion of the covariates beyond graduates' teacher evaluation ratings do not explain substantially more variation in APs' evaluation ratings relative to the number of additional parameters that are estimated in the model.

Despite the models in Table III.11 controlling for prior performance as teachers, the coefficient on the SLPP quality scores are remarkably similar to the models that do not control for the prior effectiveness of APs reported in Table III.10. The inclusion of graduates' evaluation ratings as teachers improves the overall fit and precision of the model. The magnitudes of the coefficients on PIIR scores for district partnerships and field experiences are relatively similar in size to the estimates in the corresponding models in Table III.10. For example, in the fully saturated model a one standard deviation increase in the field experience PIIR score for an SLPP is associated with a .24 points decrease in an AP's average evaluation ratings in their first two years. This estimate is slightly smaller compared to the equivalent model in Table III.10. The coefficient on district partnerships in the fully saturated model suggests that a one standard deviation increase in the partnerships PIIR score is associated with a .34 SD increase in an AP's evaluation ratings.

The patterns for SLPP scores in curriculum and selection show a slightly different pattern in the models that include graduates' average evaluation ratings as teachers from the models in Table III.10. Whereas the coefficients on selection scores were smaller in absolute magnitude and not significant in the models that do not include a control for APs' prior performance as teachers, the estimates on the selection PIIR scores in the models that control for prior performance of APs are much larger in absolute magnitude and significant in most of the models. A one standard deviation increase in an SLPP's selection quality is associated with a .18 point

decrease in the evaluation rating of APs in the fully saturated model. It is also interesting that in the model with no controls reported in column 4 of Table III.10, the coefficient on selection scores is -.021 and the coefficient in the model with no other covariates except for the rating of the graduate as a teacher is -.15. These results seem to suggest that some of the relationship between selection scores of an SLPP and APs' evaluation ratings is explained by the capacity of SLPP graduates to lead as captured in their evaluation ratings as teachers. The direction of the change in coefficients also fits with expectations because much of the positive relationship between SLPP selection quality and graduates' performance should be explained in the ability of SLPPs to select individuals with the capacity to be effective in their positions. Removing variation in the relationship between SLPP selection and graduates' performance related to the graduates' pre-existing capacity for leadership would result in a less positive relationship between SLPP selection scores and graduates' performance.

The overall results reported in Table III.11 suggest that the relationship between SLPP quality and early career performance as an AP is not purely a function of selection into SLPPs. If the relationship between SLPP quality and performance as an AP was entirely a function of selection into SLPPs, there would not be an independent relationship between prior performance metrics and performance as an AP when both SLPP quality and prior performance are in the model. The results from these analyses also suggest that quality of selection criteria and field experience requirements have a negative relationship with performance as an AP when controlling for their prior performance as a teacher. The only SLPP quality score that is consistently positively related to AP evaluation ratings is the quality of SLPPs' relationships with their district partners.

The results in Tables III.10 and III.11 can be a little difficult to parse considering that the models control for multiple factors related to SLPP quality and some of the measures of SLPP quality have negative relationships with performance as an AP. It is important to note that these results represent the difference in APs' evaluation ratings associated with marginal differences in SLPP features. In other words, the results do not represent the causal relationship between changes to SLPP features and graduates' performance. This model does not estimate what would happen to graduates' performance if an SLPP underwent a change in their programming. What these estimates represent are average differences among graduates in their performance that attend different SLPPs. An illustrative example based on the results in Table III.11 may help to clarify the interpretation of these results, and I present this example in Figure III.3. Take the case of four theoretical aspiring leaders who attend three different SLPPs. Assume all else is equal about these aspiring leaders, including their personal and school characteristics and only their SLPPs are different. Since these results are based on the model in column 6 of Table III.11, these aspiring leaders will have the same average evaluation ratings as teachers. Aspiring leader A attends SLPP W, and SLPP W has average PIIR scores in curriculum, field experiences, and district partnerships but a selection score that is one standard deviation higher than average. Aspiring leader A is predicted to have an average evaluation score of 3.42 upon graduating from SLPP W and becoming an AP. Assume aspiring leader B attends SLLP X, which has average PIIR scores in all other domains but has a score in curriculum that is one standard deviation higher than average. Aspiring leader B would have a predicted evaluation rating as an AP of 3.52. Aspiring leader C attends SLPP Y, which has average PIIR scores in all other domains but has a score in field experiences that is one standard deviation higher than average. According to the models estimated in this study, aspiring leader C would have a predicted evaluation rating of

3.36. Aspiring leader D attends SLPP Z, which has average PIIR scores in all other domain scores but has a score in district partnerships that is one standard deviation higher than average. Aspiring leader D has a predicted AP evaluation rating of 3.94 based on the fully saturated model that includes average evaluation ratings as a teacher. This illustrative example suggests that aspiring leader D is likely to have the highest evaluation ratings from attending SLPP Z.

III.6 Discussion

The results from this study suggest that there is some promise to using a document-based approach to creating measures of SLPP quality. The PIIR scores created in this study do differentiate programs from each other as demonstrated in the descriptive analyses. Along with the descriptive analyses, the results from the second research question suggest that the measures of SLPP quality are related to graduates' perceptions of their program quality. Graduates' perceptions have been used as a measure of SLPP quality in prior research (Davis et al, 2005; Levine, 2005) and the results seem to fit with the patterns described in prior research. SLPP quality in the areas of curriculum and field experience are positively associated with graduates' perceptions of their programs (Ni et al., 2019; Orr & Orphanos, 2011). This pattern provides some evidence of the PIIR scores validity as a measure of SLPP quality. Moreover, the change in the relationship between SLPP selection scores and graduates' evaluation ratings as APs when their evaluation ratings as teachers are included in the model suggests that the selection criteria is correlated to a proxy for aspiring leaders' capacity for school leadership.

The results from the regression analyses suggest that the design features of SLPPs analyzed in this study are distinct from one another and have distinct relationships with graduates' perceptions of program quality and their evaluation ratings as APs. Although the

measures of program quality are related to one another, the overall quality of an SLPP seems to have less explanatory power than the quality of specific program features. The patterns in the relationships between SLPP features and the two measures of program effectiveness assessed in this study can be difficult to parse. The features that have a positive relationship with graduates' perceptions of their SLPPs (curriculum and field experiences) are negatively related to graduates' evaluation ratings as APs. The feature that is negatively related to graduates' perceptions of SLPP quality (district partnerships) is the only feature that has a positive relationship with evaluation ratings as APs. These results suggest that perceptions of program quality and graduates' performance in leadership roles are measuring distinctly different constructs and the relationship between these measures are not particularly strong. SLPPs' graduates' average evaluation ratings as APs and their perceptions of program quality have correlations ranging from .03 to .004. Although this evidence does not suggest that graduates' perceptions of program quality is unimportant, it does suggest that measures of graduates' performance in leadership roles are also important outcomes for SLPPs to consider in evaluating and improving their programs. The results from this study seem to suggest that SLPP features can be credibly connected to graduates' performance as school leaders, just as prior studies have found that SLPPs can be credibly connected to graduates' performance in leadership roles (Gates et al., 2014; Grissom, Mitani, et al., 2019).

It is important to note that while this study does not estimate the causal relationships between SLPP features and performance as a school leader, it does suggest that any particular program features may have a more direct relationship to certain graduates' outcomes as compared to others. The example described in Figure III.3 suggests that given a theoretical choice between four programs that excel in one of the four domains of program quality included

in this study, an aspiring leader who is interested in higher evaluation ratings would be best served by attending a program that has strong district partnerships. However, this illustration necessarily simplifies the relationship between program quality features and graduates' performance as APs. In practice, SLPPs that have higher scores in one domain also tend have higher scores in the other domains, especially at the tails of the distribution of program quality. The relative importance of district partnerships actually magnifies the interconnectedness between these different domains of program quality. Programs with higher scores in district partnerships have higher scores because they include district partners in designing all of the other features of the SLPP. Moreover, aspiring leaders do not have complete authority over the SLPP they choose to attend. SLPPs also select aspiring leaders and simply because an aspiring leader wants to attend a particular SLPP does not mean that they can attend that SLPP. Further research is needed to understand how these complex relationships between program features work together to prepare aspiring leaders for school leadership.

III.7 Limitations

A critical limitation to this study is its external validity. The sample of early-career APs who are included in the study sample may not be like early-career APs across the country or even other early-career APs in Tennessee. Specifically, the results of this analysis do not address any of the long-term impacts that SLPPs features have on graduates, and the results will not be generalizable to APs who take longer than three years to start as APs. There is still a considerable percentage of APs who do not gain their first AP position until four or more years after graduation, as is suggested by the mean number of years after graduation that SLPP graduates become APs. The study findings are most applicable to the SLPPs and APs who are

most similar to those included in the study sample as is shown in Table III.2. These differences suggest that the sample of APs in this study may be different from new APs not included in the sample in ways that are unobservable as well as observed. However, APs in the sample work in schools that are relatively similar to the schools of new APs not in the sample. Additionally, Tennessee has relatively few SLPPs who are monitored regularly through the accreditation process conducted by TSBE, making it a relatively unique policy context. The quality of SLPPs may be much more variable in states with more SLPPs and less restrictive standards. Despite these limitations of the sample, this study accomplishes the modest goal of describing the relationship between common quality SLPP features and the effectiveness ratings of APs in Tennessee. Furthermore, this study provides some evidence on the usefulness of the PIIR approach to creating measures of SLPP quality based on program documents.

The second potential limitation to this study is that despite my efforts to account for all the potential confounders of the relationship between SLPP quality features and the effectiveness ratings of APs, there may still be other unobserved factors that may bias my results. One example of the type of factor that may bias my results is the quality of peers at SLPPs and interactions with peers in an SLPP. Since Ni et. al (2019) find evidence to suggest that the quality of peer relationships can influence the learning outcomes of SLPP graduates, it may be that peer relationships are related to other quality features like district partnerships and effectiveness ratings, especially for SLPPs that have strong relationships with district partnerships for selection into programs. Unfortunately, the program websites and the accreditation documents do not provide consistent information about how SLPPs structure the interactions between their students. Peer relationships is an example of the type of data that a survey like INSPIRE may capture more accurately.

A third limitation to this study is that there may be error in my measures of program quality. If the documents SLPPs submitted to TSBE for re-accreditation and program websites do not accurately reflect what is actually offered in programs, then the PIIR quality measures may be noisy. Moreover, the measures of SLPP quality may have more measurement error if programs change over the time of the study. The measures of program's quality features are based on a single moment in time and programs could change their features. There are two pieces of evidence in the literature that suggest that SLPPs are not likely to have changed substantively during the study period. The first is that some studies suggest that SLPPs have not changed much in recent years (Hess & Kelly, 2007; Levine, 2005). Second, research suggests SLPPs, when they do make changes, are responding to changes in state level policy for licensure and preparation (E. Anderson et al., 2018; Young, 2013). Since the policy in Tennessee did not change during the study period, and the re-accreditation process did not occur again until 2019, I believe it is a reasonable assumption that the features of SLPPs that graduates in 2015 experienced are likely to be similar to the features that graduates in 2017 experienced.

III.8 Conclusion

The evidence on SLPP effectiveness is limited in general and no prior study has connected SLPPs quality features to graduates' outcomes as APs. There are more APs than principals in Tennessee, and most SLPP graduates become APs before they become principals. This study highlights how using evaluation ratings of APs can be a useful tool for studying SLPPs and SLPP quality. Evaluation ratings of APs also have the potential to be used for state agencies that hold SLPPs accountable for their outcomes. This study points to why performance

metrics may be an important outcome to examine when evaluating SLPPs alongside graduates' perceptions.

Perhaps the most substantive contribution of this study is in providing evidence that suggests the novel measure of SLPP quality used in this study has potential for use in research on SLPPs. This methodology may not provide as rich a data set on programs as more substantive surveys like the INSPIRE suite, but it does produce measures that differentiate programs from one another while relying on data that may be less intensive to collect. Assuming that programs submit documents to state agencies regularly for re-authorization and accreditation, it would be relatively straight forward to code the documents of SLPPs for their features and replicate the methods in this chapter. The use of this method for measuring SLPP quality may allow for more research on the features of SLPPs, which is sorely needed in the field. While there have been more studies of SLPPs in recent years, there are still relatively few empirical studies of the relationship between SLPPs, their features, and their graduates' outcomes.

Chapter III Figures

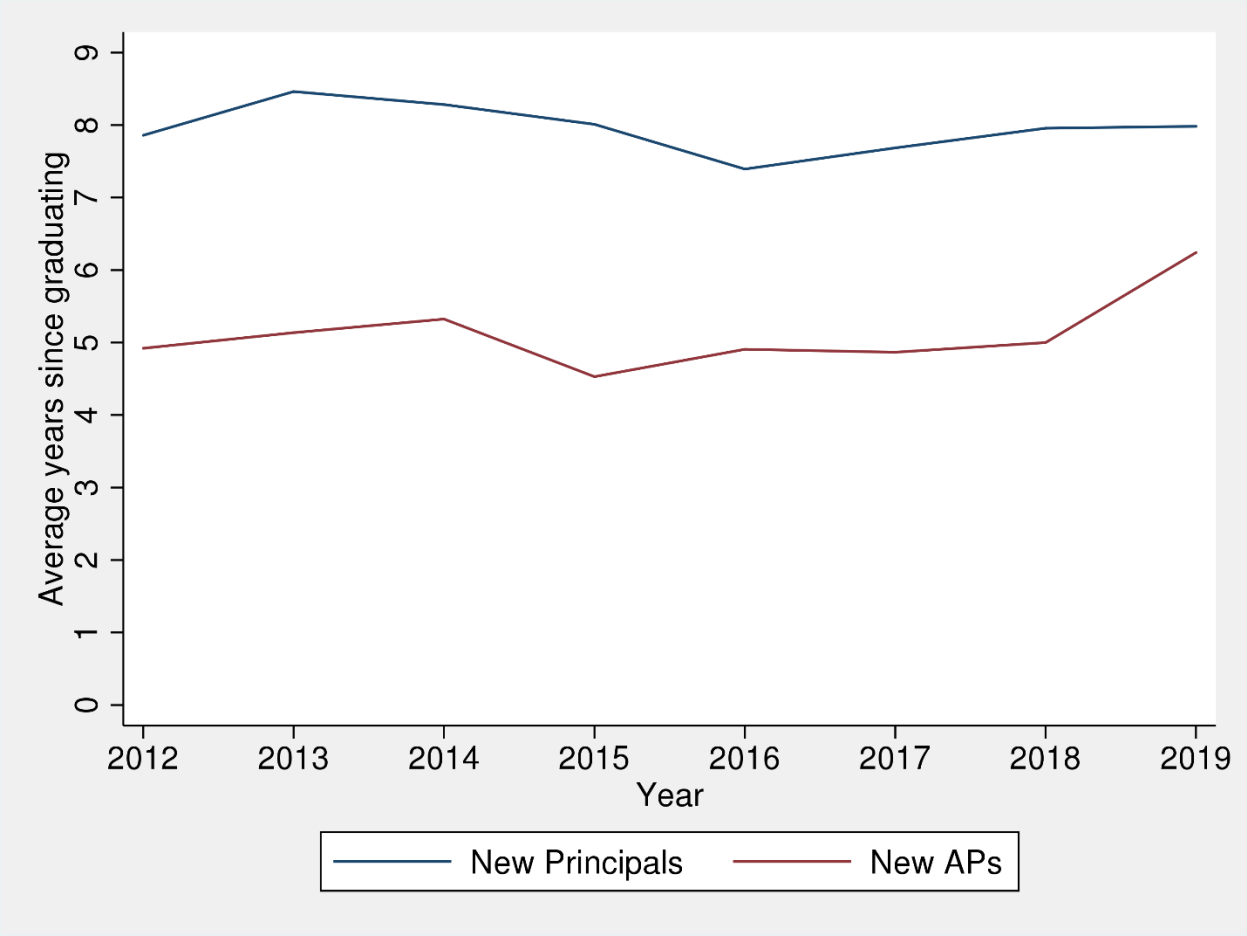


Figure III.1: New APs' and new principals' average years since graduation
Notes: New APs and new principals are defined as principals that are observed as having their first year of experience in that role each year.

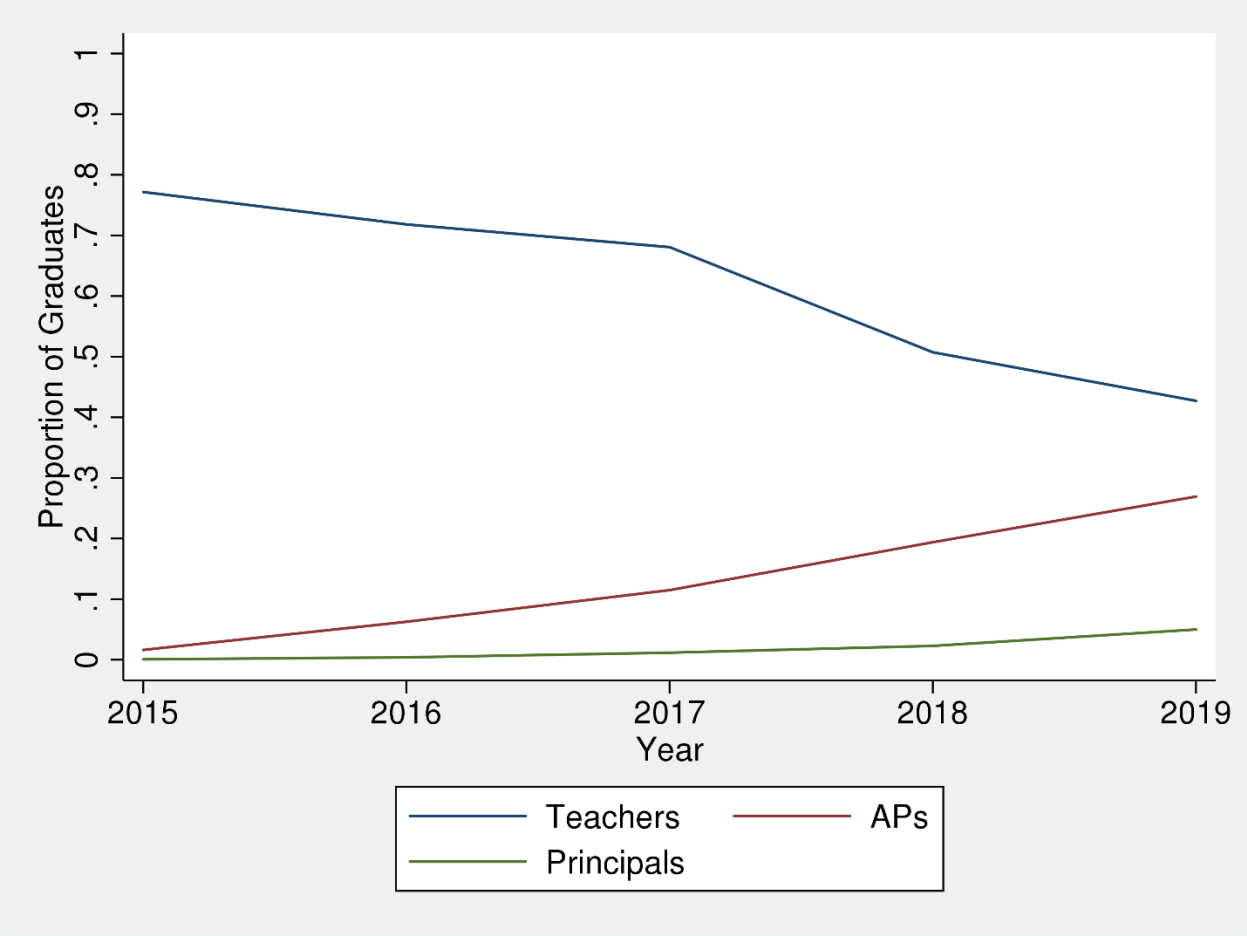
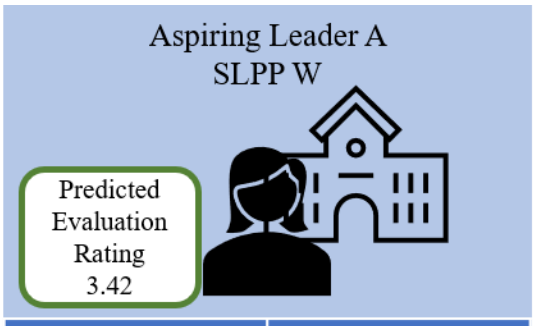
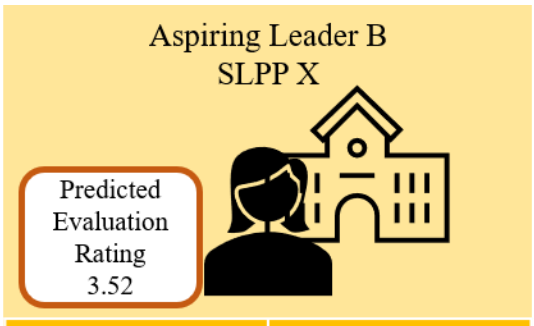


Figure III.2: Proportion of SLPP graduates in education positions

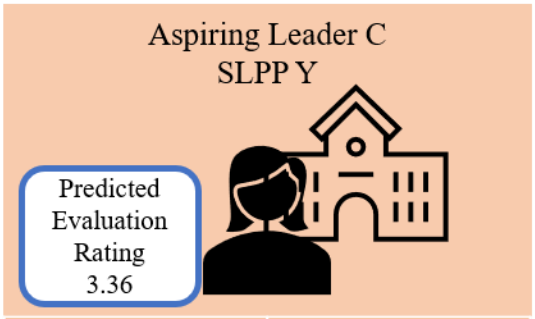
Notes: This figure presents the proportion of graduates in each cohort in the positions in the figure. These proportions may not sum to 1 because the figure does not include SLPP graduates who were in support staff roles or working in central office.



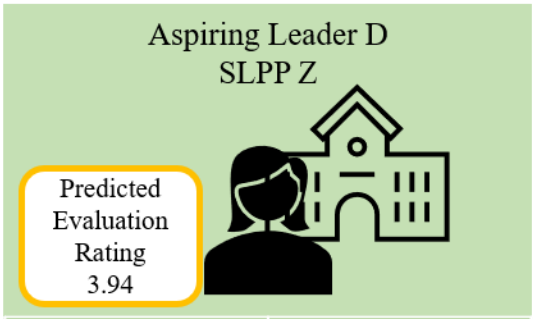
Program Feature	Feature Score
Selection	1 SD Higher
Curriculum	Average
Field Experiences	Average
District Partnerships	Average



Program Feature	Feature Score
Selection	Average
Curriculum	1 SD Higher
Field Experiences	Average
District Partnerships	Average



Program Feature	Feature Score
Selection	Average
Curriculum	Average
Field Experiences	1 SD Higher
District Partnerships	Average



Program Feature	Feature Score
Selection	Average
Curriculum	Average
Field Experiences	Average
District Partnerships	1 SD Higher

Figure III.3: Predicted evaluation ratings as an AP for four graduates from hypothetical SLPPs

Notes: Program features refer to the quality feature domains coded in this study. Feature scores refer to the PIIR scores for the given feature. Average is a PIIR score of 0 because the scores are standardized to have a mean of zero and a standard deviation of 1.

Chapter III Tables

Table III.1: Counts of SLPP graduates by year

Institution		2012	2013	2014	2015	2016	2017	Total
	Program A	6	9	25	12	12	10	74
	Program B	22	65	110	99	87	127	510
	Program C	7	5	3	14	42	66	137
	Program D	6	10	14	13	13	5	61
	Program E	1	6	3	2	14	16	42
	Program F	8	4	9	7	9	6	43
	Program G	11	18	7	6	4	8	54
	Program H	4	2	4	1	0	1	12
	Program I	73	49	69	78	66	102	437
	Program J	21	49	43	53	38	18	222
	Program K	3	13	0	1	1	18	36
	Program L	41	50	30	31	28	34	214
	Program M	71	31	50	19	18	9	198
	Program N	50	45	11	18	2	13	139
	Program O	5	8	5	13	5	6	42
	Program P	6	16	15	17	14	9	77
	Program Q	5	22	16	20	16	5	84
	Program R	43	32	31	27	26	18	177
	Program S	0	0	0	0	0	1	1
	Out of State	126	2	4	4	6	2	144
Total		509	436	449	435	401	474	2,704

Note: Program S had too few graduates identified during the study period, so it has been dropped from all other analyses.

Table III.2: Personal characteristics of sample APs as compared to other new APs

	Non-Sample	Sample
Female	0.65	0.60*
Black	0.21	0.21
Age	41.79	39.74***
Years as an AP	0.38	0.39
Years as a Teacher	0.48	0.42+
Highest Degree Master's	13.20	11.00***
Highest Degree Doctorate or Specialist	3.68	3.65
Years Since Graduation	6.45	1.74***
N	928	306

Note: * .05, **.01, ***.001

Table III.3: School characteristics of sample APs as compared to other new APs

	Non-Sample	Sample
Proportion Black Students	0.32	0.29+
Proportion Latinx Students	0.11	0.11
Proportion FRPL eligible	0.14	0.15*
Proportion with IEPs	0.61	0.60
ADM	772	819+
Achievement Index	-0.05	-0.06
Elementary	0.33	0.33
Middle	0.25	0.24
High	0.29	0.30
Urban	0.37	0.35
Suburban	0.16	0.19
Town	0.17	0.13+
Rural	0.30	0.33
N	1146	389

Note: + $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table III.4: Characteristics of SLPP graduates by cohort

	2015	2016	2017
Female	0.73	0.72	0.70
Black	0.20	0.21	0.18
Age	38.10	37.84	37.60
Highest Degree Master's	0.57	0.52	0.48**
Highest Degree Doctorate or Specialist	0.31	0.23**	0.15***
Years as a Teacher	8.95	9.34	8.51
Working as an AP	0.04	0.04	0.04
Observations	446	405	481

Note: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table III.5: Characteristics of APs by program

	Female	Black	Age	Highest: Master's	Highest: Doctorate or Specialist	Years as a Teacher
Program A	0.73	0.00	36.00	0.55	0.18	9.64
Program B	0.81	0.35	39.12	0.65	0.22	9.86
Program C	0.55	0.03	37.85	0.24	0.41	8.00
Program E	0.57	0.21	34.57	0.36	0.29	7.57
Program F	0.60	0.20	36.20	0.40	0.40	9.20
Program G	0.00	0.00	29.00	1.00	0.00	3.00
Program H	0.63	0.00	36.69	0.48	0.30	9.78
Program I	0.58	0.06	37.38	0.55	0.33	8.42
Program J	0.59	0.11	34.46	0.44	0.22	6.04
Program K	0.91	0.36	42.73	0.64	0.36	13.27
Program L	0.67	0.08	41.25	0.50	0.33	9.33
Program M	0.71	0.00	42.71	0.57	0.14	9.43
Program N	0.50	0.38	41.29	0.25	0.62	5.88
Program O	0.83	0.25	38.42	0.42	0.25	10.92
Program P	0.82	0.06	37.94	0.41	0.24	10.00
Program Q	0.83	0.08	36.92	0.50	0.25	9.75
Program R	0.82	0.55	40.64	0.36	0.55	11.64
Program S	0.75	0.58	40.58	0.25	0.25	8.00

Table III.6: Characteristics of schools employing SLPP graduates

	Proportion Black Students	Proportion Latinx Students	Proportion FRPL eligible	Proportion with IEPs	ADM	Achievement Index
Program A	0.22	0.11	0.52	0.15	897	0.46
Program B	0.40	0.09	0.64	0.14	869	0.10
Program C	0.21	0.12	0.57	0.15	846	-0.19
Program E	0.28	0.14	0.53	0.14	895	0.06
Program F	0.19	0.05	0.64	0.15	1138	0.05
Program G	0.04	0.04	0.36	0.08	1483	0.98
Program H	0.05	0.10	0.55	0.15	759	0.24
Program I	0.15	0.10	0.57	0.14	882	0.12
Program J	0.35	0.13	0.56	0.15	877	-0.19
Program K	0.35	0.08	0.63	0.14	811	-0.16
Program L	0.20	0.07	0.54	0.14	893	0.03
Program M	0.07	0.03	0.71	0.31	543	0.07
Program N	0.26	0.12	0.64	0.14	1079	0.11
Program O	0.28	0.16	0.56	0.13	898	-0.16
Program P	0.27	0.09	0.58	0.20	641	0.07
Program Q	0.21	0.08	0.44	0.15	956	-0.13
Program R	0.60	0.08	0.70	0.15	548	-0.02
Program S	0.50	0.07	0.63	0.14	966	0.18

Table III.6 cont.

	Elementary	Middle	High	Urban	Suburban	Town	Rural	Graduates
Program A	0.36	0.18	0.36	0.55	0.09	0.00	0.36	11
Program B	0.45	0.19	0.25	0.33	0.21	0.15	0.31	78
Program C	0.21	0.34	0.28	0.38	0.17	0.07	0.38	29
Program E	0.29	0.21	0.43	0.21	0.29	0.21	0.29	14
Program F	0.00	0.00	0.60	0.40	0.20	0.00	0.40	5
Program G	0.00	0.00	1.00	0.00	0.00	0.00	1.00	1
Program H	0.22	0.41	0.30	0.15	0.30	0.11	0.44	27
Program I	0.30	0.24	0.36	0.30	0.27	0.15	0.27	33
Program J	0.26	0.30	0.41	0.41	0.30	0.04	0.26	27
Program K	0.40	0.10	0.40	0.27	0.18	0.18	0.36	11
Program L	0.17	0.25	0.42	0.25	0.33	0.17	0.25	12
Program M	0.14	0.14	0.29	0.14	0.00	0.00	0.86	7
Program N	0.00	0.25	0.50	0.38	0.00	0.12	0.50	8
Program O	0.42	0.08	0.33	0.50	0.00	0.00	0.50	12
Program P	0.41	0.29	0.24	0.59	0.18	0.00	0.24	17
Program Q	0.50	0.17	0.33	0.50	0.33	0.00	0.17	12
Program R	0.45	0.36	0.00	0.36	0.27	0.36	0.00	11
Program S	0.27	0.36	0.36	0.50	0.08	0.25	0.17	12

Table III.7: PIIR Program Quality Scores

	Selection	Curriculum	Partnerships	Field Exp.	Overall Quality
Program F	-1.123	-2.332	-1.310	0.000	-1.373
Program C	-1.584	-0.287	-1.310	-1.311	-1.157
Program B	-0.163	-1.865	-0.463	-0.273	-1.036
Program I	-1.197	-0.230	-1.310	-1.311	-0.980
Program P	-1.320	-0.230	-1.310	0.000	-0.863
Program K	-0.710	-0.287	-0.463	-0.273	-0.650
Program Q	-0.396	-0.596	0.286	0.140	-0.365
Program D	0.051	-0.882	0.977	2.239	-0.343
Program M	-0.146	-0.337	0.286	0.000	-0.264
Program L	-0.504	0.311	0.286	0.000	-0.253
Program E	-0.088	0.311	-1.310	-1.311	-0.252
Program N	0.400	0.604	0.286	0.000	0.380
Program H	0.072	0.604	0.286	0.140	0.394
Program J	-0.105	0.896	0.286	0.140	0.495
Program S	0.906	0.311	0.977	2.239	0.559
Program O	1.450	0.399	-0.463	-1.334	0.658
Program R	0.997	0.311	0.977	0.913	0.740
Program G	1.730	1.805	1.646	0.000	2.090
Program A	1.730	1.493	1.646	0.000	2.219

Note: PIIR program quality scores are standardized to have a mean of zero and a standard deviation of 1. Programs are sorted by overall quality scores.

Table III.8: Coding for estimating selection PIIR score

	Selection Score	Undergrad GPA	Min Under GPA 2.7	Min Under GPA 3	GRE	Min GRE	MAT	Min MAT Score	Min MAT score 370	Min MAT score 380	Interview	Evaluation Data	Teaching License	Teaching Experience	Min 3 Yrs Experience	Recommendation	Min 2 Recs	Min 3 Recs	CO Letter
Program C	-1.584	1			1	1							1						
Program P	-1.320										1	1	1			1	1		
Program I	-1.197						1				1		1	1	1	1			
Program F	-1.123				1						1		1			1	1	1	
Program K	-0.710				1						1	1	1			1	1	1	
Program L	-0.504				1		1				1	1	1	1		1	1	1	
Program Q	-0.396	1	1								1		1			1	1	1	1
Program B	-0.163	1			1	1	1	1			1		1	1	1	1	1		
Program M	-0.146	1	1		1	1	1	1			1	1	1	1	1	1			
Program J	-0.105				1		1				1	1	1	1	1	1	1	1	
Program E	-0.088	1	1								1	1	1	1	1				1
Program D	0.051										1	1	1	1	1	1	1	1	
Program H	0.072	1	1		1	1	1	1			1					1	1	1	
Program N	0.400	1			1	1	1	1	1			1	1	1	1				
Program S	0.906	1	1	1							1	1	1	1	1	1	1	1	1
Program R	0.997	1	1	1	1	1					1	1	1	1	1	1	1		1
Program O	1.450				1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Program A	1.730	1	1		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Program G	1.730	1	1	1	1	1	1	1	1	1	1	1		1	1	1	1	1	
Counts of SLPPs		11	8	3	13	9	10	7	4	3	17	13	17	13	12	16	14	11	5

Note: A 1 represents the presence of a code in the accreditation documentation for an SLPP. An empty cell means that the SLPP did not report having that feature in the accreditation documentation. The last row provides the number of SLPPs that have that feature.

Table III.9: Relationship between APs' perceptions of SLPP quality and PIIR scores

	Rigorous		Take Again		Average	
	(1)	(2)	(3)	(4)	(5)	(6)
Overall PIIR	-0.010		0.104		0.046	
	(0.080)		(0.105)		(0.086)	
Selection		0.087		0.122		0.121
		(0.097)		(0.099)		(0.077)
Curriculum		0.234***		0.268***		0.247***
		(0.048)		(0.051)		(0.047)
Partnerships		-0.472**		-0.344*		-0.409**
		(0.153)		(0.155)		(0.134)
Field Experiences		0.257*		0.149		0.181+
		(0.096)		(0.126)		(0.099)
Adj. R-sq	-0.05	0.03	-0.04	-0.02	-0.04	-0.02
N	112	112	112	112	112	112

Notes: Standard errors clustered by SLPP in parentheses. Column headers denote dependent variables. Attend again is a likert item that asked early career APs how strongly they agree with the statement that they would attend their SLPP again given the option. Rigorous is a likert item that asked early career APs how strongly they agree with the statement that they found their SLPP to be rigorous. All models include controls for APs' personal characteristics, school characteristics, and graduating cohort fixed effects.

+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table III.10: Relationship between APs' evaluation ratings and PIIR scores

	(1)	(2)	(3)	(4)	(5)	(6)
Overall PIIR	-0.068	-0.068	-0.038			
	(0.074)	(0.071)	(0.072)			
Selection				-0.021	-0.076	-0.068
				(0.055)	(0.068)	(0.072)
Curriculum				-0.138***	-0.120*	-0.108*
				(0.032)	(0.046)	(0.047)
Partnerships				0.234*	0.285*	0.285*
				(0.105)	(0.103)	(0.124)
Field Experiences				-0.263**	-0.293***	-0.275***
				(0.074)	(0.054)	(0.059)
AP Controls		Yes	Yes		Yes	Yes
School Controls			Yes			Yes
Adj. R-sq	0.01	0.05	0.10	0.07	0.12	0.17
N	175	140	137	175	140	137

Notes: Standard errors clustered by SLPP in parentheses. Dependent variable is the average AP evaluation rating for APs in their first two years as an AP. All models include graduating cohort fixed effects.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table III.11: Relationship between APs' evaluation ratings and PIIR scores controlling for prior performance

	(1)	(2)	(3)	(4)	(5)	(6)
Overall PIIR	-0.047	-0.051	-0.030			
	(0.079)	(0.068)	(0.069)			
Selection				-0.148*	-0.160	-0.179*
				(0.068)	(0.094)	(0.082)
Curriculum				-0.117**	-0.098*	-0.073
				(0.033)	(0.040)	(0.042)
Partnerships				0.351**	0.311**	0.339**
				(0.096)	(0.090)	(0.102)
Field Experiences				-0.273***	-0.216**	-0.237**
				(0.062)	(0.055)	(0.062)
Average Rating as a Teacher	0.459***	0.324*	0.296*	0.451***	0.341*	0.320*
	(0.089)	(0.138)	(0.119)	(0.085)	(0.126)	(0.115)
AP Controls		Yes	Yes		Yes	Yes
School Controls			Yes			Yes
Adj. R-sq	0.12	0.10	0.10	0.18	0.14	0.17
N	142	114	111	142	114	111

Notes: Standard errors clustered by SLPP in parentheses. Dependent variable is the average AP evaluation rating for APs in their first two years as an AP. All models include graduating cohort fixed effects.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Chapter III Appendix Figures

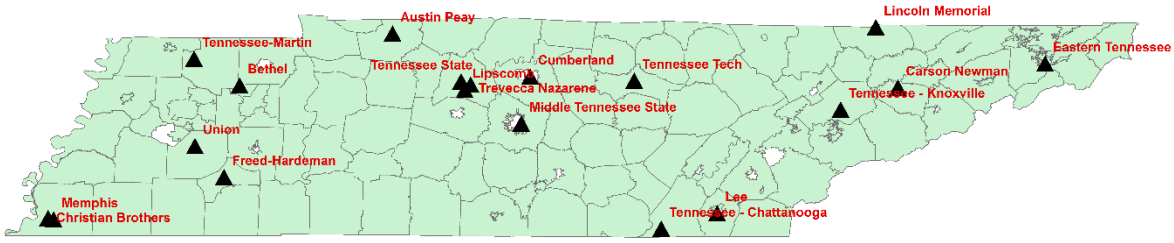


Figure III.A1 Map of SLPP main campuses in Tennessee

Chapter III Appendix Tables

Table III.A1: Coding for estimating curriculum PIIR scores

	Curriculum Score	Offers an Ed.S.	Minimum 8 Courses	Minimum 10 courses	Minimum 12 courses	Technology	Research Methods	Professional Learning	Special Education	Management and	Finances and Budgeting	Interpersonal Relationships	Community Relations	Diversity and Equity	Religion	Capstone	Action Research
Program F	-2.332																
Program B	-1.865										1						
Program D	-0.882		1							1						1	1
Program Q	-0.596		1	1			1			1	1						
Program M	-0.337	1	1	1		1				1			1	1		1	1
Program C	-0.287	1	1	1				1		1	1						
Program K	-0.287		1	1			1			1	1		1				
Program I	-0.230		1				1		1	1	1						1
Program P	-0.230	1	1				1			1			1	1			
Program E	0.311	1	1	1	1			1			1		1			1	
Program L	0.311	1	1	1			1			1	1		1			1	
Program R	0.311		1	1			1		1	1	1		1				1
Program S	0.311	1	1	1		1			1	1	1		1				
Program O	0.399	1	1				1			1	1			1	1	1	
Program H	0.604	1	1	1	1		1			1	1	1		1			
Program N	0.604	1				1	1		1	1							
Program J	0.896	1	1	1	1		1			1	1	1	1			1	
Program A	1.493	1	1	1		1	1		1	1	1		1	1		1	1
Program G	1.805	1	1	1	1	1	1			1	1	1	1		1	1	1
Counts of SLPPs		12	16	12	4	5	12	2	5	16	14	3	10	5	2	8	6

Note: A 1 represents the presence of a code in the accreditation documentation for an SLPP. An empty cell means that the SLPP did not report having that feature in the accreditation documentation. The last row provides the number of SLPPs that have that feature.

Chapter III Appendix Tables

Table III.A2: Coding for estimating Partnerships PIIR scores

	Partnerships Score	Selection	Curriculum	Faculty	Mentors	Field Placements
Program C	-1.310					
Program E	-1.310					
Program F	-1.310					
Program I	-1.310					
Program P	-1.310					
Program B	-0.463	1				
Program K	-0.463	1				
Program O	-0.463	1				
Program H	0.286	1			1	
Program J	0.286				1	1
Program L	0.286			1	1	
Program M	0.286	1	1			
Program N	0.286	1			1	
Program Q	0.286	1		1		
Program D	0.977	1	1			1
Program R	0.977	1			1	1
Program S	0.977	1		1		1
Program A	1.646	1	1	1		1
Program G	1.646	1	1		1	1
Counts of SLPPs		12	16	12	4	5

Note: A 1 represents the presence of a code in the accreditation documentation for an SLPP. An empty cell means that the SLPP did not report having that feature in the accreditation documentation. The last row provides the number of SLPPs that have that feature.

Table III.A3: Coding for estimating field experience PIIR scores

	Field Experience Scores	Field experiences earn credit	Minimum 10 experiences	Minimum 500 hours	Mentor selection criteria	Requires a mentor	Practicum is an internship	Requires exposure to diverse students
Program O	-1.334	1			1	1		
Program C	-1.311	1				1		
Program E	-1.311	1				1		
Program I	-1.311	1				1		
Program B	-0.273	1				1		
Program K	-0.273	1				1		
Program M	0.000	1						
Program N	0.000	1						
Program P	0.000	1					1	
Program F	0.000	1		1			1	1
Program L	0.000					1		
Program G	0.000				1	1	1	1
Program A	0.000	1	1		1	1	1	1
Program H	0.140	1			1	1		
Program J	0.140	1			1	1		
Program Q	0.140	1			1	1		
Program R	0.913	1			1	1		
Program D	2.239	1				1		
Program S	2.239	1				1		
Counts of SLPPs		17	1	1	7	15	4	3

Note: A 1 represents the presence of a code in the accreditation documentation for an SLPP. An empty cell means that the SLPP did not report having that feature in the accreditation documentation. The last row provides the number of SLPPs that have that feature.

Table III.A4: Inter-rater reliability

SLPP Feature	Average Kappa	Average Agreement	Items
Selection	0.213	0.65	32
Curriculum	0.311	0.765	13
Partner	0.226	0.632	5
Practicum	0.345	0.695	14
All	0.264	0.683	64

Table III.A5: Relationship between APs' perceptions of SLPP quality and PIIR scores

Panel A: Rigorous	(1)	(2)	(3)	(4)	(5)	(6)
Overall PIIR	-0.061	-0.051	-0.010			
	(0.077)	(0.088)	(0.080)			
Selection				0.019	0.084	0.087
				(0.142)	(0.136)	(0.097)
Curriculum				0.174***	0.237***	0.234***
				(0.040)	(0.051)	(0.048)
Partnerships				-0.398*	-0.511*	-0.472**
				(0.181)	(0.184)	(0.153)
Field Experiences				0.226*	0.239*	0.257*
				(0.106)	(0.105)	(0.096)
AP Controls		Yes	Yes		Yes	Yes
School Controls			Yes			Yes
adj. R-sq	-0.01	-0.07	-0.05	0.05	0.04	0.03
N	144	116	112	144	116	112
Panel B: Again	(1)	(2)	(3)	(4)	(5)	(6)
Overall PIIR	0.039	0.070	0.104			
	(0.071)	(0.096)	(0.105)			
Selection				0.015	0.115	0.122
				(0.091)	(0.100)	(0.099)
Curriculum				0.224***	0.286***	0.268***
				(0.049)	(0.060)	(0.051)
Partnerships				-0.295*	-0.404*	-0.344*
				(0.126)	(0.153)	(0.155)
Field Experiences				0.185*	0.176+	0.149
				(0.075)	(0.100)	(0.126)
AP Controls		Yes	Yes		Yes	Yes
School Controls			Yes			Yes
adj. R-sq	-0.01	-0.04	-0.04	0.02	0.01	-0.02
N	144	116	112	144	116	112
Panel C: Average	(1)	(2)	(3)	(4)	(5)	(6)
Overall PIIR	-0.008	0.009	0.046			
	(0.071)	(0.085)	(0.086)			
Selection				0.029	0.112	0.121
				(0.110)	(0.094)	(0.077)
Curriculum				0.204***	0.263***	0.247***
				(0.046)	(0.057)	(0.047)
Partnerships				-0.358*	-0.467**	-0.409**
				(0.144)	(0.149)	(0.134)
Field Experiences				0.204*	0.202*	0.181+
				(0.083)	(0.076)	(0.099)
AP Controls		Yes	Yes		Yes	Yes
School Controls			Yes			Yes
Adj. R-sq	-0.01	-0.04	-0.04	0.02	0.01	-0.02
N	144	116	112	144	116	112

Notes: Standard errors clustered by SLPP in parentheses. Column headers denote dependent variables. Attend again is a likert item that asked early career APs how strongly they agree with the statement that they would attend their SLPP again given the option. Rigorous is a likert item that asked early career APs how strongly they agree with the statement that they found their SLPP to be rigorous. All models include graduating cohort fixed effects.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table III.A6: Relationship between APs' evaluation ratings and PIIR scores (first year)

	(1)	(2)	(3)	(4)	(5)	(6)
Overall PIIR	-0.080	-0.082	-0.050			
	(0.066)	(0.054)	(0.056)			
Selection				-0.004	-0.063	-0.064
				(0.057)	(0.057)	(0.071)
Curriculum				-0.137***	-0.112**	-0.109*
				(0.029)	(0.034)	(0.039)
Partnerships				0.185	0.215*	0.237+
				(0.111)	(0.095)	(0.120)
Field Experiences				-0.221**	-0.220***	-0.205**
				(0.073)	(0.053)	(0.056)
AP Controls		Yes	Yes		Yes	Yes
School Controls			Yes			Yes
Adj. R-sq	0.00	0.08	0.12	0.04	0.11	0.14
N	165	130	127	165	130	127

Notes: Standard errors clustered by SLPP in parentheses. Dependent variable is the average AP evaluation rating for APs in their first year as an AP.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

Table III.A7: Relationship between APs' evaluation ratings and PIIR scores controlling for prior performance (first year)

	(1)	(2)	(3)	(4)	(5)	(6)
Overall PIIR	-0.051	-0.055	-0.032			
	(0.071)	(0.056)	(0.056)			
Selection				-0.123+	-0.123	-0.153+
				(0.059)	(0.076)	(0.073)
Curriculum				-0.120**	-0.100**	-0.098*
				(0.033)	(0.033)	(0.044)
Partnerships				0.323**	0.249**	0.322**
				(0.101)	(0.081)	(0.092)
Field Experiences				-0.256**	-0.161*	-0.188**
				(0.069)	(0.057)	(0.055)
Average Rating as a Teacher	0.516***	0.342*	0.315+	0.505***	0.359*	0.362*
	(0.082)	(0.149)	(0.155)	(0.081)	(0.144)	(0.159)
AP Controls		Yes	Yes		Yes	Yes
School Controls			Yes			Yes
Adj. R-sq	0.15	0.14	0.11	0.19	0.15	0.15
N	133	105	102	133	105	102

Notes: Standard errors clustered by SLPP in parentheses. Dependent variable is the average AP evaluation rating for APs in their first year as an AP.
+ p<0.10, * p<0.05, ** p<0.01, *** p<0.001

REFERENCES

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). *When Should You Adjust Standard Errors for Clustering?* (Working Paper No. 24003; Working Paper Series). National Bureau of Economic Research.

Allen, J. G., & Weaver, R. L. (2014). Learning to Lead: The Professional Development Needs of Assistant Principals. *Education Leadership Review*, *15*(2), 14–32.

Anderson, E., & Reynolds, A. (2015). The State of State Policies for Principal Preparation Program Approval and Candidate Licensure. *Journal of Research on Leadership Education*, *10*(3), 193-221.

Anderson, E., Winn, K. M., Young, M. D., Groth, C., Korach, S., Pounder, D., & Rorrer, A. K. (2018). Examining University Leadership Preparation: An Analysis of Program Attributes and Practices. *Journal of Research on Leadership Education*, *13*(4), 375–397.

Anderson, S. E., Moore, S., & Sun, J. (2009). Positioning the principals in patterns of school leadership distribution. In K. Leithwood, B. Mascall, & T. Strauss (Eds.), *Distributed leadership according to the evidence* (pp. 129–154). Routledge.

Astor, R. A., Guerra, N., & Van Acker, R. (2010). How Can We Improve School Safety Research? *Educational Researcher*, *39*(1), 69–78.

Austin, D. B., & Brown, H. L. (1970). *Report of the Assistant Principalship. Volume 3: The Study of the Secondary School Principalship.*

Austin, D. B., & Brown Jr, H. L. (1970). *Report of the Assistant Principalship. Volume 3: The Study of the Secondary School Principalship.*

Barakat, M., Reames, E., & Kensler, L. A. W. (2019). Leadership Preparation Programs: Preparing Culturally Competent Educational Leaders. *Journal of Research on Leadership Education*, *14*(3), 212–235.

Barnett, B. G., Shoho, A. R., & Oleszewski, A. M. (2012). The Job Realities of Beginning and Experienced Assistant Principals. *Leadership and Policy in Schools*, *11*(1), 92–128.

Bartanen, B. (2020). Principal Quality and Student Attendance. *Educational Researcher, 49*(2), 101-113.

Bastian, K. C., & Henry, G. T. (2015). The Apprentice Pathways to the Principalship and Student Achievement. *Educational Administration Quarterly*, *51*(4), 600–639.

Berliner, D. C. (1990). What's All the Fuss About Instructional Time. In M. Ben-Peretz & R. Bromme (Eds.), *The Nature of Time in Schools: Theoretical Concepts, Practitioner Perceptions* (pp. 3–35). Teachers College Press.

Blatchford, P., Bassett, P., Brown, P., Martin, C., Russell, A., Webster, R., & Haywood, N. (2006). *The deployment and impact of support staff in schools (DISS): Report on findings from a national questionnaire survey of schools, support staff and teachers (strand 1, wave 1, 2004)*. Institute of Education, University of London.

Bold, T., Filmer, D., Martin, G., Molina, E., Rockmore, C., Stacy, B., Svensson, J., & Wane, W. (2017). *What do teachers know and do? Does it matter? Evidence from primary schools in Africa*. The World Bank.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3–15.

Bowers, A. J., & White, B. R. (2014). Do principal preparation and teacher qualifications influence different types of school growth trajectories in Illinois? *Journal of Educational Administration*, *52*(5), 705–736.

Boyd, D., Grossman, P. L., Hammerness, K., Lankford, R. H., Loeb, S., McDonald, M., Reininger, M., Ronfeldt, M., & Wyckoff, J. (2008). Surveying the Landscape of Teacher Education in New York City: Constrained Variation and the Challenge of Innovation. *Educational Evaluation and Policy Analysis*, *30*(4), 319–343.

Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher Preparation and Student Achievement. *Educational Evaluation and Policy Analysis*, *31*(4), 416–440.

Bradshaw, C. P., Waasdorp, T. E., Debnam, K. J., & Johnson, S. L. (2014). Measuring School Climate in High Schools: A Focus on Safety, Engagement, and the Environment. *Journal of School Health*, *84*(9), 593–604.

Bryk, A. S., & Schneider, B. (2003). Trust in Schools: A Core Resource for School Reform. *Educational Leadership*, *60*(6), 40–45.

Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q., & Luppescu, S. (2010). *Organizing schools for improvement: Lessons from Chicago*. University of Chicago Press.

Burkhauser, S. (2017). How Much Do School Principals Matter When It Comes to Teacher Working Conditions? *Educational Evaluation and Policy Analysis*, *39*(1), 126–145.

Busch, S. D., MacNeil, A. J., & Baraniuk, M. S. (2012). Critical advice from practicing assistant principals for assistant principal preparation programs. In A. R. Shoho, B. Barnett, & A. K. Tooms (Eds.), *Examining the assistant principalship: New puzzles and perennial challenges for the 21st century* (pp. 35–57). Information Age Publishing, Inc.

Bush, T., & Glover, D. (2012). Distributed leadership in action: Leading high-performing leadership teams in English schools. *School Leadership & Management, 32*(1), 21–36.

Camburn, E. M., Spillane, J. P., & Sebastian, J. (2010). Assessing the Utility of a Daily Log for Measuring Principal Leadership Practice. *Educational Administration Quarterly*, *46*(5), 707–737.

Cameron, A. C., & Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, *50*(2), 317–372.

Cannata, M., Rubin, M., Goldring, E. B., Grissom, J. A., Neumerski, C. M., Drake, T. A., & Schuermann, P. (2017). Using Teacher Effectiveness Data for Information-Rich Hiring. *Educational Administration Quarterly*, *53*(2), 180–222.

Carpenter, B. W., Bukoski, B. E., Berry, M., & Mitchell, A. M. (2017a). Examining the Social Justice Identity of Assistant Principals in Persistently Low-Achieving Schools. *Urban Education*, *52*(3), 287–315.

Carpenter, B. W., Bukoski, B. E., Berry, M., & Mitchell, A. M. (2017b). Examining the Social Justice Identity of Assistant Principals in Persistently Low-Achieving Schools. *Urban Education*, *52*(3), 287–315.

Chen, K.-L., Blendinger, J., & McGrath, V. (2000). *Job Satisfaction among High School Assistant Principals*. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Bowling Green, KY.

Clark, D., Martorell, P., & Rockoff, J. (2009). School Principals and School Performance. Working Paper 38. *National Center for Analysis of Longitudinal Data in Education Research*.

Clayton, J. K., & Goodwin, M. (2015). Culturally Competent Leadership through Empowering Relationships: A Case Study of Two Assistant Principals. *Education Leadership Review*, *16*(2), 131–144.

Clifford, M., Larsen, E., Lemke, M., Chambers, D., & Swanlund, A. (2016a). *Developing Leaders: The importance and the challenges of evaluating principal preparation programs*. George W. Bush Institute and American Institutes for Research.

Clifford, M., Larsen, E., Lemke, M., Chambers, D., & Swanlund, A. (2016b). *Following the Leaders: An analysis of grauate effectiveness from five principal preparation programs*. George W. Bush Institute and American Institutes for Research.

Cochran-Smith, M. (2004). Stayers, Leavers, Lovers, and Dreamers Insights about Teacher Retention. *Journal of Teacher Education*, *55*(5), 387–392.

Cochran-Smith, M., & Villegas, A. M. (2015). Framing Teacher Preparation Research: An Overview of the Field, Part 1. *Journal of Teacher Education*, *66*(1), 7–20.

Cochran-Smith, M., Villegas, A. M., Abrams, L., Chavez-Moreno, L., Mills, T., & Stern, R. (2015). Critiquing Teacher Preparation Research: An Overview of the Field, Part II. *Journal of Teacher Education*, *66*(2), 109–121.

Cochran-Smith, M., Villegas, A. M., Abrams, L. W., Chavez-Moreno, L. C., Mills, T., & Stern, R. (2016). Research on Teacher Preparation: Charting the Landscape of a Sprawling Field. In D. H. Gitomer, C. A. Bell, D. Gitomer, & C. Bell (Eds.), *Handbook of Research on Teaching* (Fifth, pp. 439–547). American Educational Research Association.

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

Cohen, Jonathan, Mccabe, E. M., Michelli, N. M., & Pickeral, T. (2009). School Climate: Research, Policy, Practice, and Teacher Education. *Teachers College Record*, *111*, 180–213.

Cohen, Julie, & Brown, M. (2016). Teaching Quality Across School Settings. *The New Educator*, *12*(2), 191–218. https://doi.org/10.1080/1547688X.2016.1156459

Collie, R. J., Shapka, J. D., & Perry, N. E. (2011). Predicting teacher commitment: The impact of school climate and social–emotional learning. *Psychology in the Schools*, *48*(10), 1034–1048.

Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social–emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational Psychology*, *104*(4), 1189–1204.

Conley, S., Shaw, S., & Glasman, N. (2007). Correlates of Job and Professional Satisfaction among Secondary School Administrators. *Journal of School Leadership*, *17*(1), 54–88.

Corcoran, S. P., 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.

Cosner, S. (2019). What Makes a Leadership Preparation Program Exemplary? *Journal of Research on Leadership Education*, *14*(1), 98–115.

Cosner, S., Tozer, S., & Smylie, M. (2012). The Ed. D. Program at the University of Illinois at Chicago: Using Continuous Improvement to Promote School Leadership Preparation. *Planning & Changing*, *43*(1/2), 127–148.

Cosner, S., Tozer, S., Zavitkovsky, P., & Whalen, S. P. (2015). Cultivating exemplary school leadership preparation at a research intensive university. *Journal of Research on Leadership Education*, *10*(1), 11–38.

Cunningham, K. M. W., VanGronigen, B. A., Tucker, P. D., & Young, M. D. (2019). Using Powerful Learning Experiences to Prepare School Leaders. *Journal of Research on Leadership Education*, *14*(1), 74–97.

Darling-Hammond, L., Holtzman, D. J., 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*(0), 42.

Darling-Hammond, L., Meyerson, D., LaPointe, M., & Orr, M. T. (2010). *Preparing principals for a changing world: Lessons from effective school leadership programs*. John Wiley & Sons.

Davis, B. W., Gooden, M. A., & Bowers, A. J. (2017). Pathways to the Principalship: An Event History Analysis of the Careers of Teachers With Principal Certification. *American Educational Research Journal*, *54*(2), 207–240.

Davis, J. (2016). *Improving University Principal Preparation Programs: Five Themes From the Field* (p. 24). The Wallace Foundation.

Davis, S. H., & Darling-Hammond, L. (2012). Innovative principal preparation programs: What works and how we know. *Planning and Changing*, *43*(1/2), 25.

Davis, S. H., Darling-Hammond, L., Lapointe, M., & Meyerson, D. (2005). *School Leadership Study: Developing successful principals*. Stanford Education Leadership Institute.

Dorion, T. A., & Reedy, M. (2018). *Principal preparation, development, and quality and effectiveness*. University Council for Educational Administration.

Duyar, I. (2010). Relationship between school facility conditions and the delivery of instruction: Evidence from a national survey of school principals. *Journal of Facilities Management*, *8*(1), 8–25.

Eckman, E. W. (2018). A Case Study of the Implementation of the Co-Principal Leadership Model. *Leadership and Policy in Schools*, *17*(2), 189–203.

Farley-Ripple, E. N., Raffel, J. A., & Welch, J. C. (2012). Administrator Career Paths and Decision Processes. *Journal of Educational Administration*, *50*(6), 788–816. ERIC.

Findlay, N. M. (2015). Discretion in Student Discipline Insight Into Elementary Principals' Decision Making. *Educational Administration Quarterly*, *51*(3), 472–507.

Firestone, W. A. (1996). Leadership: Roles or functions? In K. Leithwood, J. Chapman, D. Corson, P. Hallinger, & A. W. Hart (Eds.), *International handbook of educational leadership and administration* (Vol. 1, pp. 395–418). Springer.

Fisher, C. W., Berliner, D. C., Filby, N. N., Marliave, R., Cahen, L. S., & Dishaw, M. M. (1981). Teaching Behaviors, Academic Learning Time, and Student Achievement: An overview. *The Journal of Classroom Interaction*, *17*(1), 2–15.

Folsom, J. S., Osborne-Lampkin, L., & Herrington, C. (2015). A Descriptive Analysis of the Principal Workforce in Florida Schools. REL 2015-068. *Regional Educational Laboratory Southeast*.

Freeston, K. R. (1987). Leader Substitutes in Educational Organizations. *Educational Administration Quarterly*, *23*(2), 45–59.

Freiberg, H. J., & Stein, T. A. (1999). Measuring, improving and sustaining healthy learning environments. In H. J. Freiberg (Ed.), *School Climate: Measuring, Improving and Sustaining Healthy Learning Environments*. Falmer Press.

Fuller, E. J., & Hollingworth, L. (2014). A Bridge Too Far? Challenges in Evaluating Principal Effectiveness. *Educational Administration Quarterly, 50*(3), 466–499. https://doi.org/10.1177/0013161X13506595

Fuller, E. J., & Hollingworth, L. (2016). Evaluating Principal-Preparation Programs Based on Placement Rates Problems and Prospects for Policymakers. *Journal of Research on Leadership Education, 11*(3) 237-271.

Fuller, E. J., & Hollingworth, L. (2018). Questioning the Use of Outcome Measures to Evaluate Principal Preparation Programs. *Leadership and Policy in Schools*, *17*(2), 167–188.

Fuller, E. J., Hollingworth, L., & An, B. P. (2016). The Impact of Personal and Program Characteristics on the Placement of School Leadership Preparation Program Graduates in School Leader Positions. *Educational Administration Quarterly*, *52*(4), 643–674.

Fuller, E. J., Hollingworth, L., & An, B. P. (2019). Exploring intersectionality and the employment of school leaders. *Journal of Educational Administration*, *57*(2), 134–151.

Fuller, E. J., Young, M. D., & Baker, B. D. (2011). Do Principal Preparation Programs Influence Student Achievement Through the Building of Teacher-Team Qualifications by the Principal? An Exploratory Analysis. *Educational Administration Quarterly*, *47*(1), 173–216.

Fuller, E. J., Young, M. D., Richardson, M. S., Pendola, A., & Winn, K. M. (2018). *The Pre-K-8 School Leader in 2018: A 10 year study*. National Association of Elementary School Principals. Retrieved from https://www.naesp.org/sites/default/files/NAESP%2010-YEAR%20REPORT_2018.pdf

Fusarelli, B. C., Fusarelli, L. D., & Drake, T. A. (2019). NC State's Principal Leadership Academies: Context, Challenges, and Promising Practices. *Journal of Research on Leadership Education*, *14*(1), 11–30.

Gajda, R., & Koliba, C. J. (2008). Evaluating and Improving the Quality of Teacher Collaboration: A Field-Tested Framework for Secondary School Leaders. *National Association of Secondary School Principals. NASSP Bulletin*, *92*(2), 133–153.

Gates, S. M., Guarino, C. M., Santibanez, L., Ghosh-Dastidar, B., Brown, A. B., & Chung, C. H. (2004). *Career paths of school administrators in North Carolina*. Rand Corporation.

Gates, S. M., Hamilton, L. S., Martorell, P., Burkhauser, S., Heaton, P., Pierson, A., Baird, M. D., Vuollo, M., Li, J. J., Lavery, D., Harvey, M., & Gu, K. (2014). *Preparing Principals to Raise Student Achievement: Implementation and effects of the New Leaders program in ten districts*. RAND corporation.

Glanz, J. (1994). Redefining the Roles and Responsibilities of Assistant Principals. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *67*(5), 283–287.

Goldhaber, D., Cowan, J., & Theobald, R. (2017). Evaluating Prospective Teachers: Testing the Predictive Validity of the edTPA. *Journal of Teacher Education*, *68*(4), 377–393.

Goldhaber, D., Holden, K., & Chen, B. (2019). *Do More Effective Teachers Become More Effective Principals?* CALDER Working Paper No. 215-0119-1.

Goldring, E. B., Grissom, J. A., Neumerski, C. M., Murphy, J., Blissett, R., & Porter, A. (2015). *Making Time for Instructional Leadership. Volume 1: Executive Summary*. Wallace Foundation.

Goldring, E. B., Grissom, J., Neumerski, C. M., Blissett, R., Murphy, J., & Porter, A. (2019). Increasing principals' time on instructional leadership: Exploring the SAM® process. *Journal of Educational Administration, ahead-of-print*(ahead-of-print).

Goldring, E. B., Huff, J., May, H., & Camburn, E. (2008). School context and individual characteristics: What influences principal practice? *Journal of Educational Administration*, *46*(3), 332–352.

Goldring, E. B., & Sims, P. (2005). Modeling Creative and Courageous School Leadership Through District-Community-University Partnerships. *Educational Policy*, *19*(1), 223–249.

Goldring, E., Grissom, J. A., Rubin, M., Neumerski, C. M., Cannata, M., Drake, T., & Schuermann, P. (2015). Make Room Value Added: Principals' Human Capital Decisions and the Emergence of Teacher Observation Data. *Educational Researcher*, *44*(2), 96–104.

Goldring, E., Huff, J., May, H., & Camburn, E. (2008). School context and individual characteristics: What influences principal practice? *Journal of Educational Administration*, *46*(3), 332–352. Education Database; ProQuest Central.

Good, R. (2008). Sharing the Secrets. *Principal Leadership*, *8*(8), 46–50.

Griffith, J. (1999). The School Leadership/School Climate Relation: Identification of School Configurations Associated with Change in Principals. *Educational Administration Quarterly*, *35*(2), 267–291.

Griffith, J. (2004). Relation of principal transformational leadership to school staff job satisfaction, staff turnover, and school performance. *Journal of Educational Administration*, *42*(3), 333–356.

Grissom, J. A. (2011). Can good principals keep teachers in disadvantaged schools? Linking principal effectiveness to teacher satisfaction and turnover in hard-to-staff environments. *Teachers College Record*, *113*(11), 2552–2585.

Grissom, J. A., & Bartanen, B. (2019). Strategic Retention: Principal Effectiveness and Teacher Turnover in Multiple-Measure Teacher Evaluation Systems. *American Educational Research Journal*, *56*(2), 514–555. https://doi.org/10.3102/0002831218797931

Grissom, J. A., Bartanen, B., & Mitani, H. (2019). Principal Sorting and the Distribution of Principal Quality. *AERA Open*, *5*(2), 2332858419850094.

Grissom, J. A., Blissett, R. S. L., & Mitani, H. (2018). Evaluating School Principals: Supervisor Ratings of Principal Practice and Principal Job Performance. *Educational Evaluation and Policy Analysis*, *40*(3), 446–472. https://doi.org/10.3102/0162373718783883

Grissom, J. A., Kalogrides, D., & Loeb, S. (2015). Using Student Test Scores to Measure Principal Performance. *Educational Evaluation and Policy Analysis*, *37*(1), 3–28.

Grissom, J. A., & 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. A., Loeb, S., & Master, B. (2013). Effective Instructional Time Use for School Leaders: Longitudinal Evidence From Observations of Principals. *Educational Researcher*, *42*(8), 433–444.

Grissom, J. A., Loeb, S., & Mitani, H. (2015). Principal time management skills: Explaining patterns in principals' time use, job stress, and perceived effectiveness. *Journal of Educational Administration*, *53*(6), 773–793.

Grissom, J. A., Mitani, H., & Woo, D. S. (2019). Principal Preparation Programs and Principal Outcomes. *Educational Administration Quarterly*, *55*(1), 73–115.

Gronn, P. (2000). Distributed Properties: A New Architecture for Leadership. *Educational Management & Administration*, *28*(3), 317–338.

Gronn, P. (2002). Distributed leadership as a unit of analysis. *The Leadership Quarterly*, *13*(4), 423–451.

Gronn, P., & Hamilton, A. (2004). 'A Bit More Life in the Leadership': Co-Principalship as Distributed Leadership Practice. *Leadership and Policy in Schools*, *3*(1), 3–35.

Grossman, P. L., Hammerness, K. M., McDonald, M., & Ronfeldt, M. (2008). Constructing Coherence: Structural Predictors of Perceptions of Coherence in NYC Teacher Education Programs. *Journal of Teacher Education*, *59*(4), 273–287.

Grubb, W. N., & Flessa, J. J. (2006). "A Job Too Big for One": Multiple Principals and Other Nontraditional Approaches to School Leadership. *Educational Administration Quarterly*, *42*(4), 518–550.

Gurley, D. K., Anast-May, L., & Lee, H. T. (2015). Developing Instructional Leaders Through Assistant Principals' Academy: A Partnership for Success. *Education and Urban Society*, *47*(2), 207–241.

Hackmann, D. G., & McCarthy, M. M. (2013). What Constitutes a Critical Mass? An Investigation of Faculty Staffing Patterns in Educational Leadership Programs. *Journal of Research on Leadership Education*, *8*(1), 5–27.

Hallinger, P., & Murphy, J. (1985). Assessing the Instructional Management Behavior of Principals. *The Elementary School Journal*, *86*(2), 217–247.

Harris, A. (2009). Distributed leadership and knowledge creation. In K. Leithwood, B. Mascall, & T. Strauss (Eds.), *Distributed leadership according to the evidence* (pp. 271–284). Routledge.

Harris, A. (2013). Distributed leadership: Implications for the role of the principal. *Journal of Management Development, 31*(1), 7–17.

Harris, A., & Spillane, J. (2008). Distributed leadership through the looking glass. *Management in Education, 22*(1), 31–34. https://doi.org/10.1177/0892020607085623

Hausman, C. S., & Goldring, E. B. (2001). Teachers' Ratings of Effective Principal Leadership: A Comparison of Magnet and Nonmagnet Elementary Schools. *Journal of School Leadership*, *11*(5), 399–423.

Hausman, C. S., Nebeker, A., McCreary, J., & Donaldson, G. (2002). The worklife of the assistant principal. *Journal of Educational Administration*.

Heck, R. H., & Hallinger, P. (2009). Assessing the Contribution of Distributed Leadership to School Improvement and Growth in Math Achievement. *American Educational Research Journal*, *46*(3), 659–689.

Heck, R. H., & Hallinger, P. (2010). Collaborative Leadership Effects on School Improvement: Integrating Unidirectional- and Reciprocal-Effects Models. *The Elementary School Journal*, *111*(2), 226–252.

Henry, G. T., Purtell, K. M., Bastian, K. C., Fortner, C. K., Thompson, C. L., Campbell, S. L., & Patterson, K. M. (2014). The Effects of Teacher Entry Portals on Student Achievement. *Journal of Teacher Education, 65*(1), 7–23.

Hess, F. M., & Kelly, A. P. (2005). An Innovative Look, a Recalcitrant Reality: The Politics of Principal Preparation Reform. *Educational Policy, 19*(1), 155–180.

Hess, F. M., & Kelly, A. P. (2007). Learning to lead: What gets taught in principal-preparation programs. *The Teachers College Record*, *109*(1), 244–274.

Hitt, D. H., & Tucker, P. D. (2016). Systematic Review of Key Leader Practices Found to Influence Student Achievement: A Unified Framework. *Review of Educational Research*, *86*(2), 531–569.

Hoffman, J. V., Wetzel, M. M., Maloch, B., Greeter, E., Taylor, L., DeJulio, S., & Vlach, S. K. (2015). What can we learn from studying the coaching interactions between cooperating teachers and preservice teachers? A literature review. *Teaching and Teacher Education*, *52*, 99–112.

Hollingworth, L., & Dude, D. (2009). Race, Gender, and School Leadership in a State with Shifting Student Demographics. *Journal of Women in Educational Leadership*, *7*(3), 175–194.

Honig, M. I., & Donaldson Walsh, E. (2019). Learning to Lead the Learning of Leaders: The Evolution of the University of Washington's Education Doctorate. *Journal of Research on Leadership Education*, *14*(1), 51–73.

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

Houchens, G., Niu, C., Zhang, J., Miller, S. K., & Norman, A. D. (2018). Do Differences in High School Principal and Assistant Principal Perceptions Predict Student Achievement Outcomes? *NASSP Bulletin*, *102*(1), 38–57.

Hoy, W. K., Hannum, J., & Tschannen-Moran, M. (1998). Organizational Climate and Student Achievement: A Parsimonious and Longitudinal View. *Journal of School Leadership*, *8*(4), 336–359.

Hoy, W. K., & Tschannen-Moran, M. (2007). The conceptualization and measurement of faculty trust in schools. *Essential Ideas for the Reform of American Schools*, 87–114.

Huang, T., Hochbein, C., & Simons, J. (2018). The relationship among school contexts, principal time use, school climate, and student achievement. *Educational Management Administration & Leadership*, *48*(2), 305-323.

Hughes, G. D. (2012). Teacher Retention: Teacher Characteristics, School Characteristics, Organizational Characteristics, and Teacher Efficacy. *The Journal of Educational Research*, *105*(4), 245–255.

Hulpia, H., Devos, G., & Van Keer, H. (2010). The Influence of Distributed Leadership on Teachers' Organizational Commitment: A Multilevel Approach. *The Journal of Educational Research*, *103*(1), 40–52.

Keesor, C. A. (2005). Administrative Visibility and Its Effect on Classroom Behavior. *NASSP Bulletin*, *89*(643), 64–73.

Kelly, C. G., & Chesser, M. C. (2019). *Delaware School Administrator Funding Analysis: Submitted to House Concurrent Resolution 39 Task Force to Study School District Consolidation*. Institute for Public Administration: Biden School of Public Policy & Administration.

Kerr, S., & Jermier, J. M. (1978). Substitutes for leadership: Their meaning and measurement. *Organizational Behavior and Human Performance*, *22*(3), 375–403.

Kinsler, J. (2011). Understanding the black–white school discipline gap. *Economics of Education Review*, *30*(6), 1370–1383.

Kmetz, J. T., & Willower, D. J. (1982). Elementary School Principals' Work Behavior. *Educational Administration Quarterly*, *18*(4), 62–78.

Korach, S., Anderson, E., Hesbol, K., Tabron, L., Candelarie, D., Kipp, P., & Miller-Brown, E. (2019). Interdependence and Reciprocity: Partnership Ethos at the University of Denver. *Journal of Research on Leadership Education*, *14*(1), 31–50.

Koru, J. M. (1993). The Assistant Principal: Crisis Manager, Custodian, or Visionary? *NASSP Bulletin*, *77*(556), 67–71.

Kwan, P., & Walker, A. (2008). Vice-principalship in Hong Kong: Aspirations, competencies, and satisfaction. *School Effectiveness and School Improvement*, *19*(1), 73–97.

Lee, T., Cornell, D., Gregory, A., & Fan, X. (2011). High Suspension Schools and Dropout Rates for Black and White Students. *Education and Treatment of Children*, *34*(2), 167–192.

Leithwood, K., & Jantzi, D. (2009). A Review of Empirical Evidence About School Size Effects: A Policy Perspective. *Review of Educational Research*, *79*(1), 464–490.

Leithwood, K., Louis, K. S., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning: A review of the evidence linking leadership to student learning*. The Wallace Foundation.

Levine, A. (2005). *Educating school leaders. New York: Education Schools Project*.

Liebowitz, D. D., & Porter, L. (2019). The Effect of Principal Behaviors on Student, Teacher, and School Outcomes: A Systematic Review and Meta-Analysis of the Empirical Literature. *Review of Educational Research*, *89*(5), 785–827.

Lochmiller, C. R., & Karnopp, J. R. (2016). The politics of coaching assistant principals: Exploring principal control. *International Journal of Mentoring and Coaching in Education; Bingley*, *5*(3), 203–220.

Lochmiller, C. R., & Mancinelli, J. L. (2019). Principals' instructional leadership under statewide teacher evaluation reform. *The International Journal of Educational Management*, *33*(4), 629–643.

MacNeil, A. J., Prater, D. L., & Busch, S. (2009). The effects of school culture and climate on student achievement. *International Journal of Leadership in Education*, *12*(1), 73–84.

Madhlangobe, L., & Gordon, S. P. (2012). Culturally Responsive Leadership in a Diverse School: A Case Study of a High School Leader. *NASSP Bulletin*, *96*(3), 177–202.

Makles, A. (2012). Stata tip 110: How to get the optimal k-means cluster solution. *Stata Journal*, *12*(2).

Marshall, C., & Hooley, R. M. (2006). *The assistant principal: Leadership choices and challenges*. Corwin Press.

May, H., Huff, J., & Goldring, E. (2012). A longitudinal study of principals' activities and student performance. *School Effectiveness and School Improvement*, *23*(4), 417–439.

Mayrowetz, D. (2008). Making Sense of Distributed Leadership: Exploring the Multiple Usages of the Concept in the Field. *Educational Administration Quarterly*, *44*(3), 424–435.

McClellan, R. L., & Casey, P. (2015). Lost in Transition? Campus Leaders' Professional Pathways. *Journal of School Leadership*, *25*(4), 720–757.

Mertz, N. T. (2006). The Organizational Socialization of Assistant Principals. *Journal of School Leadership*, *16*(6), 644–675.

Metzker, B. (2003). *Time and Learning. ERIC Digest*. ERIC Clearinghouse on Educational Management.

Militello, M., Fusarelli, B. C., Mattingly, A., & Warren, T. (2015). “We Do What We’re Told”: How Current Assistant Principals Practice Leadership and how They Wish They Could. *Journal of School Leadership*, *25*(2), 194–222.

Mitchell, R. M., Kensler, L., & Tschannen-Moran, M. (2018). Student trust in teachers and student perceptions of safety: Positive predictors of student identification with school. *International Journal of Leadership in Education*, *21*(2), 135–154.

Moore, D. C. (2013). Race specialists: What a black administrator ought to be and do. *Journal of School Leadership*, *23*(3), 490-510.

Morgan, T. L. (2018). Assistant Principals’ Perceptions of the Principalship. *International Journal of Education Policy and Leadership*, *13*(10).

Munoz, M. A., & Barber, H. M. (2011). Assistant Principals in High-Stakes Accountability Environments: The Effects of Job Attributes and School Characteristics. *Educational Assessment, Evaluation and Accountability*, *23*(2), 131–142.

Murphy, J., Smylie, M., Mayrowetz, D., & Louis, K. S. (2009). The role of the principal in fostering the development of distributed leadership. *School Leadership & Management*, *29*(2), 181–214.

Murphy, J., & Vriesenga, M. (2006). Research on school leadership preparation in the United States: An analysis. *School Leadership & Management*, *26*(2), 183–195.

Neumerski, C. M. (2013). Rethinking Instructional Leadership, a Review: What Do We Know About Principal, Teacher, and Coach Instructional Leadership, and Where Should We Go From Here? *Educational Administration Quarterly*, *49*(2), 310–347.

Neumerski, C. M., Grissom, J. A., Goldring, E. B., Drake, T. A., Rubin, M., Cannata, M., & Schuermann, P. (2018). Restructuring Instructional Leadership: How Multiple-Measure

Teacher Evaluation Systems Are Redefining the Role of the School Principal. *The Elementary School Journal*, *119*(2), 270–297.

Ni, Y., Hollingworth, L., Rorrer, A., & Pounder, D. (2017). The Evaluation of Educational Leadership Preparation Programs. In M. D. Young & G. M. Crow (Eds.), *Handbook of Resaearch on the Education of School Leaders* (2nd ed., pp. 285–307). Routledge.

Ni, Y., Rorrer, A. K., Pounder, D., Young, M. D., & Korach, S. (2019). Leadership matters: Preparation program quality and learning outcomes. *Journal of Educational Administration*, *57*(2), 185–206.

Noell, G. H., Burns, J. M., & Gansle, K. A. (2019). Linking Student Achievement to Teacher Preparation: Emergent Challenges in Implementing Value Added Assessment. *Journal of Teacher Education*, *70*(2), 128–138.

Oleszewski, A., Shoho, A. R., & Barnett, B. (2012). The development of assistant principals: A literature review. *Journal of Educational Administration*, *50*(3), 264–286.

Oliver, R. (2003). Assistant Principal Job Satisfaction and Desire To Become Principals. *Education Leadership Review*, *4*(2), 38–46.

Oliver, R. (2005). Assistant Principal Professional Growth and Development: A Matter that Cannot Be Left to Chance. *Educational Leadership and Administration: Teaching and Program Development*, *17*, 89–100.

Orr, M. T. (2010). Pipeline to Preparation to Advancement: Graduates' Experiences In, Through, and Beyond Leadership Preparation. *Educational Administration Quarterly*, *47*(1), 114–172.

Orr, M. T., & Orphanos, S. (2011). How Graduate-Level Preparation Influences the Effectiveness of School Leaders: A Comparison of the Outcomes of Exemplary and

Conventional Leadership Preparation Programs for Principals. *Educational Administration Quarterly*, *47*(1), 18–70.

Parylo, O., Zepeda, S. J., & Bengtson, E. (2013). The different faces of principal mentorship. *International Journal of Mentoring and Coaching in Education*, *1*(2), 120–135.

Pellicer, L. O., Anderson, L. W., Keefe, J. W., Kelly, E. A., & McCleary, L. E. (1988). *High School Leaders and Their Schools. Volume I: A National Profile*. National Association of Secondary School Principals.

Petrides, L., Jimes, C., & Karaglani, A. (2014). Assistant principal leadership development: A narrative capture study. *Journal of Educational Administration*, *52*(2), 173–192.

Pitner, N. J. (1986). Substitutes for Principal Leader Behavior: An Exploratory Study. *Educational Administration Quarterly*, *22*(2), 23–42.

Pitner, N. J. (1988). Leadership Substitutes: Their Factorial Validity in Educational Organizations. *Educational and Psychological Measurement*, *48*(2), 307–315.

Podsakoff, P. M., & MacKenzie, S. B. (1997). Kerr and Jermier's substitutes for leadership model: Background, empirical assessment, and suggestions for future research. *The Leadership Quarterly*, *8*(2), 117–132.

Polanchek, C. (2019). *Basic Education Program*. Tennessee Comptroller of the Treasury: Research and Education Accountability. Retrieved from https://comptroller.tn.gov/office-functions/research-and-education-accountability/legislative-toolkit/bep.html

Portin, B. S., Knapp, M. S., Dareff, S., Feldman, S., Russell, F. A., Samuelson, C., & Yeh, T. L. (2009). *Leadership for Learning Improvement in Urban Schools* (p. 141). Center for the Study of Teaching and Policy.

Portin, B. S., Schneider, P., DeArmond, M., & Gundlach, L. (2003). *Making Sense of Leading Schools: A Study of the School Principalship* (p. 62). The Center on Reinventing Public Education.

Pounder, D. (2012). School Leadership Preparation and Practice Survey Instruments and Their Uses. *Journal of Research on Leadership Education, 7*(2), 254–274.

Powell, D. (2017). *Inference with Correlated Clusters.* RAND Corporation.

Preston, J. P., & Barnes, K. E. R. (2017). Successful Leadership in Rural Schools: Cultivating Collaboration. *The Rural Educator, 38*(1), 6–15.

Purkey, S. C., & Smith, M. S. (1983). Effective Schools: A Review. *The Elementary School Journal*, *83*(4), 427–452.

Reardon, S. F., & Raudenbush, S. W. (2006). A Partial Independence Item Response Model for Surveys with Filter Questions. *Sociological Methodology*, *36*(1), 257–300.

Ringel, J., Gates, S., Chung, C., Brown, A., & Ghosh-Dastidar, B. (2004). *Career paths of school administrators in Illinois: Insights from an analysis of state data*.

Rocha, R. R., & Hawes, D. P. (2009). Racial Diversity, Representative Bureaucracy, and Equity in Multiracial School Districts*. *Social Science Quarterly*, *90*(2), 326–344.

Rocque, M. (2010). Office Discipline and Student Behavior: Does Race Matter? *American Journal of Education*, *116*(4), 557–581.

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.

Ronfeldt, M., Brockman, S. L., & Campbell, S. L. (2018). Does Cooperating Teachers' Instructional Effectiveness Improve Preservice Teachers' Future Performance? *Educational Researcher*, *47*(7), 405–418.

Ronfeldt, M., & Campbell, S. L. (2016). Evaluating Teacher Preparation Using Graduates' Observational Ratings. *Educational Evaluation and Policy Analysis*, *38*(4), 603–625.

Ronfeldt, M., Reininger, M., & Kwok, A. (2013). Recruitment or Preparation? Investigating the Effects of Teacher Characteristics and Student Teaching. *Journal of Teacher Education*, *64*(4), 319–337.

Schoen, L. T., & Teddlie, C. (2008). A new model of school culture: A response to a call for conceptual clarity. *School Effectiveness and School Improvement*, *19*(2), 129–153.

Schweig, J., Hamilton, L., & Baker, G. (2019). *School and Classroom Climate Measures: Considerations for Use by State and Local Education Leaders*. RAND Corporation.

Searby, L., Browne-Ferrigno, T., & Wang, C. (2017). Assistant Principals: Their Readiness as Instructional Leaders. *Leadership and Policy in Schools*, *16*(3), 397–430.

Searby, L. J., Wang, C., & Browne-Ferrigno, T. (2015, August 4). *The Profile of the "Ready" Assistant Principal as Instructional Leader in Alabama*. National Council of Professors of Educational Administration Annual Conference, Arlington, VA.

Sebastian, J., Allensworth, E., Wiedermann, W., Hochbein, C., & Cunningham, M. (2018). Principal Leadership and School Performance: An Examination of Instructional Leadership and Organizational Management. *Leadership and Policy in Schools*, *0*(0), 1–23.

Skiba, R. J., Chung, C.-G., Trachok, M., Baker, T. L., Sheya, A., & Hughes, R. L. (2014). Parsing Disciplinary Disproportionality Contributions of Infraction, Student, and School

Characteristics to Out-of-School Suspension and Expulsion. *American Educational Research Journal*, *51*(4), 640–670.

Smylie, M. A. (1988). The Enhancement Function of Staff Development: Organizational and Psychological Antecedents to Individual Teacher Change. *American Educational Research Journal*, *25*(1), 1–30.

Smylie, M. A., Mayrowetz, D., Murphy, J., & Louis, K. S. (2007). Trust and the Development of Distributed Leadership. *Journal of School Leadership*, *17*(4), 469–503.

Sorensen, L. C., Bushway, S. D., & Gifford, E. J. (2020). *Getting Tough? The Effects of Discretionary Principal Discipline on Student Outcomes. (EdWorkingPaper: 20-216)*. Annenberg Institute at Brown University.

Spear-Swerling, L., & Zibulsky, J. (2014). Making time for literacy: Teacher knowledge and time allocation in instructional planning. *Reading and Writing*, *27*(8), 1353–1378.

Spillane, J. P. (2005). Distributed Leadership. *The Educational Forum*, *69*(2), 143–150.

Spillane, J. P., Camburn, E. M., & Pareja, A. S. (2007). Taking a Distributed Perspective to the School Principal's Workday. *Leadership and Policy in Schools*, *6*(1), 103–125.

Spillane, J. P., Halverson, R., & Diamond, J. B. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher*, *30*(3), 23–28.

Springer, M. G., Swain, W. A., & Rodriguez, L. A. (2016). Effective Teacher Retention Bonuses: Evidence From Tennessee. *Educational Evaluation and Policy Analysis*, *38*(2), 199–221.

Stallings, J. (1980). Allocated Academic Learning Time Revisited, or beyond Time on Task. *Educational Researcher*, *9*(11), 11–16.

Strunk, K. O. (2011). Are Teachers' Unions Really to Blame? Collective Bargaining Agreements and Their Relationships with District Resource Allocation and Student Performance in California. *Education Finance and Policy*, *6*(3), 354–398.

Strunk, K. O., Cowen, J. M., Goldhaber, D., Marianno, B. D., Kilbride, T., & Theobald, R. (2018). It is in the Contract: How the Policies Set in Teachers' Unions' Collective Bargaining Agreements Vary Across States and Districts. *Educational Policy*, *32*(2), 280–312.

Strunk, K. O., & Grissom, J. A. (2010). Do Strong Unions Shape District Policies?: Collective Bargaining, Teacher Contract Restrictiveness, and the Political Power of Teachers' Unions. *Educational Evaluation and Policy Analysis*, *32*(3), 389–406.

Struyven, K., Dochy, F., & Janssens, S. (2010). 'Teach as you preach': The effects of student-centered versus lecture-based teaching on student teachers' approaches to teaching. *European Journal of Teacher Education*, *33*(1), 43–64.

Sun, A. (2012). Sharing Instructional Leadership as Assistant Principals in an Accountability-Oriented Environment. In A. R. Shoho, B. G. Barnett, & A. K. Tooms (Eds.), *Examining the Assistant Principalship: New puzzles and perennial challenges for the 21st century* (pp. 153–180). Information Age Publishing, Inc.

Sun, A., & Shoho, A. R. (2017). Assistant Principals' Perceptions of Value Added to School Success. *Journal of School Leadership*, *27*(4), 456-490.

Tennessee Department of Education (TDOE). (2014). *Tennessee Basic Education Program 2.0: Handbook for computation*. Tennessee Department of Education.

Tennessee State Board of Education. (2019). *Tennessee Basic Education Program—BEP Blue Book*. Tennessee State Board of Education.

Thapa, A., Cohen, J., Guffey, S., & Higgins-D'Alessandro, A. (2013). A Review of School Climate Research. *Review of Educational Research, 83*(3), 357–385.

Tian, M., Risku, M., & Collin, K. (2016). A meta-analysis of distributed leadership from 2002 to 2013: Theory development, empirical evidence and future research focus. *Educational Management Administration & Leadership, 44*(1), 146–164.

Tschannen-Moran, M. (2014). *Trust Matters: Leadership for Successful Schools*. John Wiley & Sons.

Tschannen-Moran, M., & Gareis, C. R. (2015). Faculty trust in the principal: An essential ingredient in high-performing schools. *Journal of Educational Administration, 53*(1), 66–92.

Turnball, B. J., Anderson, L. M., Riley, D. L., MacFarlane, J. R., & Aladjem, D. K. (2016). *Building a Stronger Principalship Vol 5: The Principal Pipeline Initiative in Action*. Policy Studies Associates, Inc.

Turnbull, B. J., Anderson, L. M., Riley, D. L., MacFarlane, J. R., & Aladejm, D. K. (2016). *Building a Stronger Principalship Vol 5: The Principal Pipeline Initiative in Action* (PR-2445-WF). Policy Associates, Inc.

Uline, C., & Tschannen-Moran, M. (2008). The walls speak: The interplay of quality facilities, school climate, and student achievement. *Journal of Educational Administration, 46*(1), 55–73.

U.S. Department of Education. (2016). *Non-Regulatory Guidance for Title II, Part A: Building Systems of Support for Excellent Teaching and Leading*. U.S. Department of Education.

Van Houtte, M. (2005). Climate or Culture? A Plea for Conceptual Clarity in School Effectiveness Research. *School Effectiveness and School Improvement, 16*(1), 71–89.

Vannest, K. J., & Hagan-Burke, S. (2010). Teacher Time Use in Special Education. *Remedial and Special Education*, *31*(2), 126–142.

Virgilio, I., Teddlie, C., & Oescher, J. (1991). Variance and Context Differences in Teaching at Differentially Effective Schools. *School Effectiveness and School Improvement*, *2*(2), 152–168.

Walker, A., & Kwan, P. (2009). Linking Professional, School, Demographic, and Motivational Factors to Desire for Principalship. *Educational Administration Quarterly*, *45*(4), 590–615.

Wang, E., Gates, S., Herman, R., Mean, M., Perera, R., Tsai, T., Whipkey, K., & Andrew, M. (2018). *Launching a Redesign of University Principal Preparation Programs: Partners Collaborate for Change*. RAND Corporation.

Wang, M.-T., & Degol, J. L. (2016). School Climate: A Review of the Construct, Measurement, and Impact on Student Outcomes. *Educational Psychology Review*, *28*(2), 315–352.

Watson, K. (2005). How Are Assistant Heads Affecting Primary School Management and How Do Their Opinions, Attitudes and Beliefs Affect Their Work? *Management in Education*, *19*(4), 24–27.

Weller, L. D., & Weller, S. J. (2002). *The assistant principal: Essentials for effective school leadership*. Corwin Press.

Whitehorn, T. (2010). *School Support Staff Topic Paper*. School Analysis and Research Division, Department for Education.

Williams, J. A., Lewis, C., Starker Glass, T., Butler, B. R., & Hoon Lim, J. (2020). The Discipline Gatekeeper: Assistant Principals' Experiences With Managing School Discipline in Urban Middle Schools. *Urban Education*, Advanced online publication.

Winn, K. M., Anderson, E., Groth, C., Korach, S., Pounder, D., Rorrer, A., & Young, M. D. (2016). *A Deeper Look: INSPIRE Data Demonstrates Quality in Educational Leadership Preparation*. UCEA.

Wolcott, H. F. (2003). *The man in the principal's office: An ethnography*. Rowman Altamira.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Nelson Education.

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.

Young, M. D. (2013). Is state-mandated redesign an effective and sustainable solution? *Journal of Research on Leadership Education*, *8*(2), 247–254.

Young, M. D. (2015a). Effective Leadership Preparation: We Know What It Looks Like and What It Can Do. *Journal of Research on Leadership Education*, *10*(1), 3–10.

Young, M. D. (2015b). The Leadership Challenge: Supporting the Learning of All Students. *Leadership and Policy in Schools*, *14*(4), 389–410.

Young, M. D., & Crow, G. M. (Eds.). (2017). *Handbook of Research on the Education of School Leaders*. Routledge.

Young, M. D., Crow, G. M., Murphy, J., & Ogawa, R. T. (2009). *Handbook of research on the education of school leaders*. Routledge.

Young, M. D., Orr, M. T., & Tucker, P. D. (2012). *Institutional and Program Quality Criteria: Guidance for master's and doctoral programs in educational leadership*. University Councile for Educational Administration.

Young, M. D., & Rorrer, A. K. (2012). Promoting the Utilization of Educational Leadership Research in Preparation, Practice, and Policy. *Journal of Research on Leadership Education*, *7*(2), 195–216.

Zellner, L., Ward, S. M., McNamara, P., Gideon, B., Camacho, S., & Edgewood, S. D. (2002, February 15). *The Loneliest Job in Town: Sculpting the Recruitment and Retention of the Principal*. Annual Meeting of the Southwest Educational Research Association, Austin, TX.

Zhu, J., Liao, Z., Yam, K. C., & Johnson, R. E. (2018). Shared leadership: A state-of-the-art review and future research agenda. *Journal of Organizational Behavior*, *39*(7), 834–852.

Zuckerman, S. J., Wilcox, K. C., Durand, F. T., Lawson, H. A., & Schiller, K. S. (2018). Drivers for Change: A Study of Distributed Leadership and Performance Adaptation During Policy Innovation Implementation. *Leadership and Policy in Schools*, *17*(4), 618–646.