Three Essays on School Reform

By

Adam Kho

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

August 10, 2018

Nashville, Tennessee

Approved:

Gary T. Henry, Ph.D.

Christopher A. Candelaria, Ph.D.

Carolyn J. Heinrich, Ph.D.

Ron Zimmer, Ph.D.

ACKNOWLEDGEMENTS

I am grateful for the Leadership, Policy, and Organizations faculty at Vanderbilt University's Peabody College, who have provided me with continued support and mentorship throughout my doctoral program. In particular, I want to thank my advisor and committee chair, Gary T. Henry, who has patiently guided me and advocated for me. I am also grateful for the tutelage of Ron Zimmer who has served as a secondary advisor and mentor to me. Many thanks to the other members of my dissertation committee, Carolyn J. Heinrich and Christopher A. Candelaria, for their helpful and thoughtful feedback, and Andrew McEachin for his continued mentorship.

Fellow students of LPO who have befriended, encouraged, and challenged me – I thank you. In particular, I am grateful for the mutual support of my cohort – Seth, Emily, Chris, Luis, and Samantha. Members of the previous cohort, the North Carolina and Tennessee longitudinal datasets team, the Tennessee Turnaround research team, the Tennessee Education Research Alliance, and my colleagues in Payne 106 have all provided great camaraderie and support. To all of you, thank you.

I also want to thank my family for their ongoing love and support – my parents who trekked halfway across the world with nothing to pursue a better life for our family, my sister and brother-in-law, and the extended Kho family. Lastly, thank you Bradford for your enduring patience and support through this dissertation process.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

INTRODUCTION

Various school reforms have been implemented to improve the quality of education for the nation's most disadvantaged students. It is important to evaluate these reforms to understand both their intended and unintended effects. This dissertation examines three recent education reforms in Tennessee – the Community Eligibility Provision, systemic effects of the Memphis Innovation Zone school turnaround initiative, and Tennessee's charter schools.

The first chapter is a study of Tennessee's Community Eligibility Provision (CEP), a program implemented in 2014-15 that changed the structure of its school meal programs to allow schools with larger percentages of economically disadvantaged students to provide free meals to all students, regardless of income level. The purpose of CEP was to ensure that all students had access to nutritious meals. In particular, the program sought to reduce the social stigma that prevented students who were eligible for free meals from participating. While previous research has examined the effects of various school breakfast and lunch programs, this chapter contributes to the literature by focusing specifically on the Community Eligibility Provision, the most recent iteration of the federal school meal program, and its impacts on various student outcomes. In addition, this chapter examines a previously unexplored element – how the revised school meal program impacts students most likely affected by the social stigma of participating in free school meals.

In chapter 2, I study the unintended consequences of recruiting high-performing teachers for school turnaround. Many districts and states have begun implementing teacher incentive programs to attract high-performing teachers to low-performing schools. Previous research has found that these programs are effective at raising student achievement at the receiving schools. However, no studies to my knowledge have examined the resulting effect at the schools losing these teachers to the incentive program. This chapter focuses on these systemic effects as a result of the Memphis Innovation Zone (iZone), one of the most successful recent turnaround initiatives in the nation. Understanding the unintended consequences of the Memphis iZone coupled with previous work that evaluated the intended effects helps to provide a more holistic picture of the impact of school turnaround initiatives.

Chapter 3 examines the effectiveness of Tennessee's charter schools. Public charter schools have been in existence since the 1990's. However, Tennessee only adopted charter schools in 2004, making it one of the latest states to accept charter schools in their public education landscape.   In addition to an evaluation of Tennessee's charter schools as a whole, this chapter tests a theory from the community development literature that argues asset-based development of local charter schools should yield more successful outcomes than needs-based development of charter schools managed by national charter management organizations. The results of this analysis are particularly salient for charter school authorizers assessing applications for future charter schools.

Using administrative data from Tennessee in conjunction with various quasi-experimental approaches, this dissertation evaluates the effects of these three education reforms. Understanding the impacts of these school reforms is essential for making policy-relevant decisions to continue or modify current school reforms and to improve the quality of Tennessee's education system. However, each chapter also provides important policy implications nationwide, as all states manage school meal programs and seek to turnaround low-performing schools, and the number of charter schools continues to expand in nearly every state.

CHAPTER 1

FREE MEALS FOR ALL: THE EFFECT OF THE COMMUNITY ELIGIBILITY PROVISION PROGRAM IN TENNESSEE

Introduction

Since its beginnings in the mid-twentieth century, access to free and reduced-price, nutritious school meals has been an important policy for improving the academic performance of low-income students. In recent years, it has received even more attention, particularly with former First Lady Michelle Obama's focus on raising a healthier generation of children. Support for this program has developed from an extant literature showing that (1) providing meals at school and (2) eating breakfast and lunch improves not only cognitive functioning and academic performance (Belot & James, 2011; Benton & Parker, 1998; Hoyland, Dye & Lawton, 2009; Kleinman et al., 2002; Meyers et al., 1989; Murphy et al., 1998; Rampersaud et al., 2005; Taras, 2005; Wesnes et al., 2003; Worobey & Worobey, 1999), but behavioral and future educational outcomes as well (Hinrichs, 2010; Pollitt & Mathews, 1998; Richter et al., 1997). These benefits may be a function, at least in part, of decreased absences and tardiness by having the availability of meal programs in schools (Belot & James, 2009; Kleinman et al., 2002; Meyers et al., 1989; Murphy et al., 1998; Rampersaud et al., 2005; Taras, 2005).

Unfortunately, not all students partake in school meals, including some of those who are eligible for free or reduced-price meals (FRPM). For some students, after factoring in the amount of time required to get their food, check out, and find a seat, there simply isn't enough time left to eat (Bartfield & Kim, 2010; Bergman et al., 2000; Cohen et al., 2015; Gordon, Crepinsek & Condon, 2007; Marples & Spillman, 1995; PerryUndem, 2013; Sanchez et al., 1999; Westervelt, 2013). For other students, the cost of school meals is too high. Research has found, however, that lower-priced meals are associated with a greater uptake (Gordon et al., 2007; Ham et al., 2002; Maurer, 1984). Food quality is also a factor. The results of several student and parent surveys have shown that many simply do not like the food provided at school, whether it be the nutritional quality, variety of foods offered, or personal preferences (Gordon et al., 2007; Marples & Spillman, 1995; PerryUndem, 2013; Zellner, 2016). Lastly, many students who are eligible for FRPM still choose not to take advantage of the program, suffering from peer pressure. To these students, eating subsidized school meals is associated with poverty, and the stigma surrounding this status is avoidable by simply not partaking (Askelson et al., 2015, 2017; Bailey-Davis et al., 2013; Bhatia, Jones & Reicker, 2011; Glantz et al., 1994; Gordon et al., 2007; Lambert et al., 2007; Marples & Spillman, 1995; McDonnell et al., 2004; Mitcherva & Powell, 2009; Poppendieck, 2010).

To address these concerns as well as others, the federal government, in 2010, passed the Healthy, Hunger-Free Kids Act (HHFKA). HHFKA provided an option called the Community Eligibility Provision (CEP) for districts and schools with large enough populations of low-income students to receive free school meals for all students, regardless of family income. This would allow schools and districts opting into the program to streamline meal service operations, providing students more time to eat. And since all students, regardless of family income status, could partake, low-income students would not feel singled out and stigmatized for waiting in the lunch line or in some cases a special line, being called up to receive a meal ticket, or reciting their free lunch ID to the cashier. In addition to these concerns, the CEP program would help reduce paperwork and administrative costs tied to the former individual application program (USDA, 2015).

As a result of providing free meals for all students and removing the stigma attached with receiving free meals due to their access only for low-income students, we might expect a greater take-up of school meals, which may in turn translate into increased academic performance, as prior research indicates that eating breakfast and/or lunch improves cognitive functioning and academic performance (Belot & James, 2009; Benton & Parker, 1998; Hoyland, Dye & Lawton, 2009; Kleinman et al., 2002; Meyers et al., 1989; Murphy et al., 1998; Rampersaud et al., 2005; Taras, 2005; Wesnes et al., 2003; Worobey & Worobey, 1999). Further, provided with guaranteed food options, students may be encouraged to attend school more frequently to gain access to free meals, thus boosting attendance. Similarly, unenrolled students might be encouraged to enroll in school to take advantage of free meals. Lastly, past research has found evidence that hunger is associated with greater misconduct (Gunderson, 2003). Providing free meals to all students may therefore reduce the number of behavior-related incidents and the discipline rates associated with the incidents.

We may, however, suspect that these effects may vary based on specific subgroups of students. For instance, students who have never qualified for FRPM (i.e., students from wealthier backgrounds) may be more likely to experience a positive effect if (1) they begin to eat the school meals and (2) those meals are healthier than their previous alternatives (i.e., meals from home, if any). Students who pre-CEP paid a reduced price for school meals may also benefit, as previous research suggests a negative relationship between meal price and participation (Gleason, 1996; Gordon, Crepinsek & Condon, 2007; Maurer, 1984). Lastly, students who were eligible to receive free meals *and* participated before the policy change may not experience a large impact after CEP took effect.

This leaves a fourth subgroup of great interest to policymakers – the students who are eligible to participate in the FRPM program but choose not to. In this paper, I tackle the effect of CEP on students affected by this stigmatization in three ways. First, while some of these students completed the eligibility application but chose not to partake, others chose not to complete the eligibility application at all. By identifying those who switched off FRPM status in years prior to CEP, I can assess the effect of CEP on this subsample of students who may have been affected by the stigma.

Second, to determine if CEP helped to reduce stigma, I examine its effects by school level. Dahl & Scholz (2011) find in a nationally representative study that students in middle school and high school may be more influenced by the social stigma of FRPM and less likely to participate in free school meals than younger children. Therefore, I assess if the effects are differentially greater in secondary schools relative to elementary schools.

Third, in order for students to experience the stigma, there must be a subpopulation of the school that is not eligible and does not participate in FRPM. Following this logic, the larger the population of FRPM-ineligible students, the greater the possible effect of stigmatization on those eligible for FRPM. Thus, by assessing how the policy change is moderated by the percentage of students eligible for FRPM in the school, I can determine to what extent CEP has been an effective solution for addressing issues of stigmatization for those eligible but not participating in FRPM.

Tennessee, the state that is the locus of this study, opted into the CEP program in 2014-15, the first year it became available nationwide. Using statewide, student-level longitudinal data from the state of Tennessee, this study seeks to answer the following research questions about Tennessee's CEP program:

1) To what extent does participation in CEP increase student attendance, on-time grade progression, achievement, and/or enrollment or decrease the number of students with discipline infractions?
2) Does the CEP program appear to have reduced the stigma associated with participating in free school meals?
    a. Are the relationships between CEP participation and student outcomes moderated by previous free-and-reduced-price-meal (FRPM) eligibility status – (a) students who were never eligible for FRPM, (b) students who were previously eligible for reduced-price meals, (c) students who were previously eligible for free meals, and (d) students who were not eligible for FRPM in the year prior to CEP but were in year(s) prior?
    b. Do the relationships between CEP participation and student outcomes vary by school level?
    c. Is the relationship between CEP participation and student outcomes moderated by the percentage of students in a school that were individually eligible for FRPM?

In 2016-17, 85% of the schools eligible for CEP in Tennessee adopted the program. However, 220 eligible schools in Tennessee (and thousands more throughout the country) still have not signed on. The results of this analysis will be important for schools, districts, and policymakers to better understand if adoption of CEP was a beneficial decision for students' educational outcomes. Further, if the results of this analysis reveal positive effects for students, state policymakers should further encourage these schools and districts to commit to the program.

The next section reviews a brief history of the National School Lunch Program and previous research surrounding the effects of free and reduced meal programs on student outcomes. Following this, I provide some more context on the CEP program in Tennessee and elaborate on the research questions of interest. I then describe the data and methods used for the analysis, followed by a discussion of the results and findings. I conclude with final thoughts for future research and policy implications.

Literature Review

National School Lunch & School Breakfast Programs

The current day school lunch program began with the National School Lunch Act in 1946. The program was created "as a measure of national security, to safeguard the health and well-being of the Nation's children…" (Gunderson, 1993). Many states, districts, and individual schools had previously created school meal programs to combat concerns of malnutrition and inability to learn due to hunger, but soon found their own programs unsustainable during the Great Recession and throughout World War II. At this point, the federal government committed to providing partial funding for states' purchase of food and equipment necessary for preparing school meals. In line with many programs already in place, the funding would require schools to serve students meals meeting a minimal nutritional requirement as determined by the federal government and that meals be offered at no or reduced costs for students from low-income families. Schools would then be reimbursed for food purchases based on the number of meals served meeting the nutritional requirements. In 1966, the National School Lunch Program (NSLP) was extended to include a pilot School Breakfast Program (SBP), which became permanent in 1975.

To determine if students qualified for free or reduced-price meals, students' family incomes were considered. Students whose family incomes fell below 130% of the federal poverty level qualified for free meals. Students whose family incomes fell between 130 and 185% of the federal poverty level paid a reduced price. And students whose family incomes were greater than 185% of the federal poverty level paid "full price." All meals, however, were subsidized by the federal government, including "full price" meals, just at a lower rate. As of the 2016-17 school year, schools in the contiguous United States could receive up to $2.04 and $3.39 for each free breakfast and lunch served, respectively; $1.74 and $2.99 for each reduced-price breakfast and lunch served, respectively; and $0.29 and $0.44 for each full-price paid breakfast or lunch served, respectively (USDA, 2016)[1].

Early evaluations of NSLP and SBP relied primarily on personal testimony: "I see definite personality changes when a child doesn't get lunch"; "Since getting free lunch she has shown a marked improvement in attitude. Last year she was a major discipline problem"; and "Attendance has improved by approximately ¾-day per student. The majority of the children have shown a good increase in weight (some 10-12 pounds)… This has also created a better home-school relationship." are three early quotes from teachers and administrators who supported the program (Gunderson, 1993). Though not directly tied to NSLP or SBP, this was substantiated by research studies from South Africa, Chile, and New York City that found continued evidence of restricted intellectual development in malnourished children (Gunderson, 1993). In 1984, the first large-scale evaluation of NSLP and SBP was published reporting that school lunches boosted the nutritional intake of students partaking as compared to other students

[1] The rates discussed are maximum reimbursement rates. Reimbursement rates can vary based on the percentage of the school that qualifies for free meals, though this amount is currently no less than $0.34 lower than the maximum. Reimbursement rates are greater for schools located in Alaska, Hawaii, and Puerto Rico.

who had alternative meals, but school breakfasts did not (Radzikowski & Gale, 1984). Meyers and colleagues (1989) soon thereafter investigated the impact of SBP on educational outcomes of elementary school students in an experiment in Lawrence, Massachusetts. The researchers conducted analyses of variance and found a positive association between school breakfast participation and test score gains and a negative association between school breakfast participation and tardiness rates. A similar experiment with a classroom of preschool students also yielded better performance on a variety of cognitive tasks when students participated in SBP (Worobey & Worobey, 1999). In a review of multiple SBP studies internationally, Taras (2005) also found a positive association between SBPs and improved attendance rates, tardiness rates, and academic performance, particularly in undernourished populations.

More recent evaluations have utilized quasi-experimental methods to identify plausibly causal impacts of school meals. Bhattacharya and colleagues (2006) used the nationally representative National Health and Nutritional Examination Survey (NHANES) III conducted between 1988 and 1994, which includes data on dietary intake and results from physical exams and laboratory tests performed by certified doctors, to investigate the effects of breakfast on nutrition. The study concluded that SBP positively impacted students' nutrition, including reduced calories from fat and increased intake of necessary nutrients and vitamins. In an analysis linking SBP to academic performance, Frisvold (2015) found that students residing in states with a mandate to provide breakfast through SBP (as of 2004) scored about 9% of a standard deviation better in math and 5 to 12% of a standard deviation better in reading on the National Assessment of Education Progress than students in states that do not hold the same mandate.

While past research has mostly found positive results on a variety of student outcomes for school breakfast programs, there have been mixed results for school lunch programs. Using a unique approach comparing siblings in which one participates in NSLP but the other does not, Dunifon & Kowaleski-Jones (2003) found no evidence of effects of school lunch programs on student behavior, parent-reported health conditions, or test scores. Hinrichs (2009) explored a number of long-term outcomes and also found no evidence of effects of NSLP on health conditions but did find that students who participate in NSLP more often experience increased educational attainment.

Universal Free Meals

Early amendments to the National School Lunch Act allowed for the expansion of its implementation in new settings and with new students. For instance, additional funding was provided to build cafeterias and to hire more staff to accommodate increasing interest in the program. Further, the meal programs have been extended to include other youth-serving institutions, such as daycares, and have since expanded to providing meals for summer learning programs and reimbursable snacks throughout the school day. However, a growing concern that has hindered student participation despite being eligible and having access to free meals has not been tackled until recently. Many students choose not to participate because of the stigma attached to receiving free meals (Bartfield & Kim, 2010; Gordon, Crepinsek & Condon, 2007; PerryUndem, 2013; Askelson et al., 2015; Askelson et al., 2016; Bailey-Davis et al., 2013; Bhatia et al., 2011; Glantz et al., 1994; Lambert et al., 2007; Mitcherva & Powell, 2009; Poppendieck, 2010). A survey with randomly selected students in Cincinnati Public Schools found that 18.5% of all students would participate more often in school meals if their friends did

(Marples & Spillman, 1995). In a study interviewing students about receiving free school meals, stigma arose as a common issue (Glantz et al., 1994). Students said:

> "You keep it to yourself if you get it for free. So many people are embarrassed to tell others that they get their lunch free. It's embarrassing because the people who are on it are the ones who don't have that much money."
>
> "Even if some kids are eligible, they may not want to take it because they think they're just too cool, or others may not want to be seen with those who do take it because they want to fit in."
>
> "Some people are just too resistant. They have too much pride to eat it. They would rather not eat lunch and be hungry than be seen eating lunch."
>
> "Hunger never wins out over embarrassment, unless you're really starving. A growling stomach where you know you'll eat at dinner, you just won't eat lunch."

Even parents consider stigma an issue, both for their children and for themselves. In Glantz and colleague's study (1994), 20% of parents said they did not even apply for FRPM for their children because of stigma-related issues. Parents responded, "they're [my kids are] not going to be taking a handout" (Glantz et al., 1994). In other studies, parents said, "we don't accept help from the government" (Poppendieck, 2010) and were afraid of being labeled as "lazy" or "irresponsible" (Askelson et al., 2015).

Stigma of participating in FRPM has certainly hindered take-up rates of school meals, but as early as the mid-1990's, many schools and districts had begun to implement a program that combatted this issue – Universal Free Meals (UFM). With UFM, all students, regardless of family income status, could partake in free school meals. Students no longer had to stand in a special line, receive a special ticket, or punch in a special ID number that indicated their low-income qualification status, making it more difficult to identify whether students came from low-income families and thus removing the associated stigma. In fact, in two different studies evaluating UFM programs in Minnesota and Milwaukee, student interviews revealed that students no longer experienced the stigma associated with the prior qualification procedure under the new UFM program (Wahlstrom & Begalle, 1999; Lent & Emerson, 2007). These two studies, in addition to several other evaluations in New York City, Philadelphia, Baltimore, and a national study, also concluded that UFM led to greater participation rates in school meals, ensuring that more students received sustenance during the school day (Leos-Urbel et al., 2013; Murphy et al., 1998; Bernstein et al., 2004; Schanzenbach & Zaki, 2014).

Several of these evaluations took the next step to examine the effect of UFM on several educational outcomes, yielding a mixed bag of results. In Minnesota, a series of interviews and an analysis of means after the program revealed increased levels of concentration, alertness, and energy in students, fewer discipline problems, and better attendance rates and test scores as the result of a pilot universal school breakfast program (Wahlstrom & Begalle, 1999). Similar results were found in Boston, Milwaukee, Philadelphia and Baltimore with improved attendance, grades, health, and behavior (Kleinman et al., 2002; Lent & Emerson, 2007; Murphy et al., 1998). However, the USDA implemented a randomized control trial titled "School Breakfast Program Pilot Project" in six school districts from 2000-01 to 2002-03. While participation did increase, there were no effects found on other educational outcomes or health outcomes (Bernstein et al., 2004). A reanalysis of this data using quasi-experimental methods yielded the same results, as did another analysis of a UFM program in New York City (Schanzenbach & Zaki, 2014; Leos-Urbel et al., 2013). In North Carolina, one school district reverted to the individual eligibility rules in 2008-09 after a year of UFM when concerns grew regarding the large increase in school meal uptake and the ability to fund so many meals. As a result, school meal participation decreased, but there were no observed effects on test scores or attendance (Ribar & Haldeman, 2013). While many perceived improvements in student outcomes, more rigorous research methods have yet to find positive impacts of UFM on student educational outcomes.

Community Eligibility Provision

Regardless of educational outcomes, there was a unanimous conclusion that students' participation in school meal programs increased as a result of UFM. Further, no studies have found any harmful effects on students' health, education, or otherwise. These findings, coupled with the original intent of the National School Lunch Act signed by Harry Truman, led to the development of the Community Eligibility Provision that accompanied the Healthy, Hunger-Free Kids Act signed by then President Barack Obama in 2010. The Community Eligibility Provision allowed districts like the one in North Carolina that could not afford to maintain their UFM program to reinstate universal free meals for all students. It could also allow many other districts to begin implementation of a UFM program so that students in the poorest schools in the country could gain access to free meals without being stigmatized based on their socioeconomic status.

Not every school, however, qualifies for the Community Eligibility Provision (CEP). CEP is reserved for the neediest communities. In order to determine eligibility for the CEP program, each district is first assigned an identified student percentage (ISP). This percentage identifies the percentage of students in the district (1) whose families are eligible for federal assistance programs such as the Supplemental Nutrition Assistance Program (SNAP) or the Temporary Assistance for Needy Families Program (TANF) or (2) who are currently in foster care or are migrant, homeless, a runaway, or enrolled in Head Start. Districts with an ISP of 40% or more can choose to switch from the individual student eligibility application to the CEP program. If the district chooses to switch to CEP, all schools within the district switch, regardless of the individual school's student population. If the district chooses not to opt in or the district does not qualify as a whole, individual schools are then assessed. Likewise, schools are assigned an ISP according to their student population and those with an ISP greater than or equal to 40% can choose to opt in to CEP. Additionally, groups of schools can band together to qualify based on a group level ISP.

Under CEP, federal reimbursement formulas for school meals also changed. Previously, schools were reimbursed per meal served differently for free meals, reduced-price meals, and full price meals. Under CEP, because all students are served free meals, there is no distinction between students of these tiers. Instead, reimbursements are determined based on the qualifying level's ISP (district, school, or group). Districts, schools, or groups of schools with ISPs of 62.5% or greater would receive 100% of meals served reimbursed at the free meals rate under CEP. Those with ISPs between 40% and 62.5% would multiply their ISP by 1.6 to determine the percentage of meals served that would be reimbursed at the free meals rate. The remaining percentage (subtracted from 100%) is then reimbursed at the full-price paid rate. For instance, if a school under CEP has an ISP of 50%, 80% of the meals served (50% x 1.6) would be reimbursed at the free meals rate, and 20% (100% - 80%) would be reimbursed at the full-price paid rate. Using the reimbursement rates mentioned prior, this school would be reimbursed $280 for 100 lunches served in one day. This is compared to the previous reimbursement system in which the same school would be reimbursed only $191.50 if the 50 students who do not qualify for free meals all pay full price; $319 if the 50 students who do not qualify for free meals all pay the reduced price (this does not take into account any payment from students). To break even between the two reimbursement calculations, about 69% of the students who do not qualify for free meals would need to qualify for reduced-price meals.

The above calculation serves as one reason why a school or district may have chosen not to switch to CEP. If a significant proportion of their students can receive school meals at a reduced price under the prior system, the district or school may actually lose funding under the CEP program. A number of researchers have investigated other reasons schools and districts have chosen not to switch to CEP. These include concerns regarding funding uncertainty, staffing issues, and the need for a reason to collect mandatory student socioeconomic data for the school or district (Logan et al., 2014; Moore, 2016).

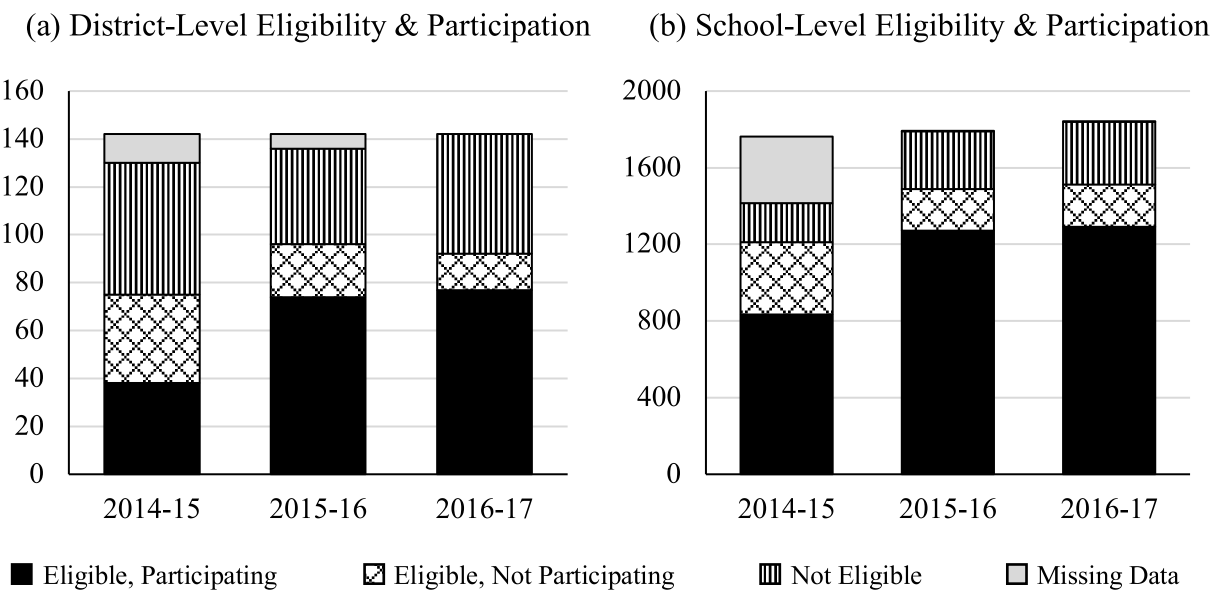
On the other hand, there are many potential benefits of CEP. First, students would have more time to eat (USDA, 2015). Research shows that many students do not have enough time to eat during school breakfast and lunch hours (Bartfield & Kim, 2010; Gordon, Crepinsek & Condon, 2007; Marples & Spillman, 1995; PerryUndem, 2013; Bergman et al., 2000; Cohen et al., 2016; Westervelt, 2013). Under the CEP program, there is no need for students to wait in a line to pay. Students simply grab their food and find a seat. (The opposite, however, could also be argued, as providing free meals for all might induce a greater demand for food, creating a longer serving line.) Second, paperwork and administrative costs are reduced (USDA, 2015). Under the previous FRPM eligibility system, staff first identified students using the direct certification procedure, matching students against a system that identifies families that already qualify for government assistance programs. Following this, schools had to collect individual student applications from those not yet in the system in order to determine which students were eligible for free or reduced-price meals based on their household income (Segel et al., 2016). Under CEP, schools rely solely on the first system, completely eliminating paper applications. Any students that would have qualified under the second step are captured by the 1.6 multiplier as previously discussed. Further, schools would not be required to assess for CEP qualification again until four years after the prior assessment. (Again, the opposite, however, could be argued. As previously indicated, some districts still required collection of student socioeconomic data, so a new form had to be created, distributed, collected, and analyzed.)

A CEP pilot program began in the 2011-12 school year, with three to four states phasing in each year. An external evaluation of the pilots found a number of additional benefits cited by pilot participants, including overall increased revenue, despite previous funding concerns; parental satisfaction, particularly for families under financial burdens; and decreased administrative burden for both school administrators and foodservice staff (Logan et al., 2014). However, the greatest benefit from CEP is ensuring that school meals are available for all students, regardless of income, and particularly for students who previously did not partake due to the associated stigma. Consequently, though not fully supported by some of the past literature, increased participation in school meals may lead to benefits in student achievement and behavior.

Tennessee Context

Tennessee, the state that is the locus of this study, joined the CEP program in 2014-15, the first year that it became available nationwide. It was one of nine non-pilot states in which over 50% of the states' eligible school districts opted in (Segel et al., 2016). Figure 1 shows the eligibility and participation rates for the first three years of implementation (a) by district and (b) by school. In its first year, 75 of Tennessee's 142 districts were eligible for CEP as a full district. However, only 38 adopted the program, or 51% of eligible districts. In year two, the percentage of districts eligible for CEP grew by 13 percentage points, and 77% of eligible school districts adopted CEP. This remained fairly stable into the following school year. At the school level, 1,212 of the state's 1,764 schools were eligible for CEP in the first year Tennessee offered CEP, but only 69% of these eligible schools opted into the CEP program. This increased dramatically in the following year, with 1,270 schools or 85% of eligible schools opting into the CEP program.

Figure 1. CEP Eligibility and Participation



This study addresses two research questions to investigate the impact of the CEP program on Tennessee's students:

RQ1. To what extent does participation in CEP increase student access to free school meals, student attendance, on-time grade progression, achievement, and/or enrollment or decrease the number of students with discipline infractions?

A number of studies have found that school and district participation in CEP has increased school meal participation rates (Logan et al., 2014; Henry, 2015; Hong, 2015). There has also been anecdotal evidence of the impact of CEP on student attendance and behavior (Neuberger, 2016; Moore, 2016). However, few studies have yet estimated the causal impact of

CEP on student educational outcomes. In this study, I utilize a comparative interrupted time series design to identify a plausibly causal impact of CEP on student attendance, on-time grade progression, achievement, enrollment, and behavioral outcomes.

First, I descriptively examine the change in student access to free and reduced-price school meals over time. I then examine the effect of increased student access to free school meals on attendance rates, testing the hypothesis that free meals will encourage students to attend schools to gain access to the free meals or that the additional nutrition reduces absences related to improper nutrition. In turn, better attendance may translate into better academic performance, which may also occur as a function of better nourishment. Enrollment rates may also increase as more students seek access to free meals. Lastly, in conjunction with prior research, students may be, as a function of being regularly nourished, better behaved resulting in fewer discipline problems.

It is important to note that take-up of free meals is not available in the administrative dataset available for this project. Therefore, the analysis described above depends on the viability of the assumption that access to free school meals increases take-up of school meals. Previous research shows that school and district participation in CEP programs have increased meal participation rates nationwide (Logan et al., 2014; Hong, 2015), which provides confidence in the assumption for this study.

RQ2. Does the CEP program appear to have reduced the stigma associated with participating in free school meals?

One of the primary reasons for implementing CEP was to reduce the stigma associated with participating in free meals, encouraging students who were affected to eat. Research question 2 examines whether this is the case. Note that without knowing exactly which students were affected by the stigma, it is impossible to precisely examine whether CEP helped to reduce stigma. However, I examine whether patterns within the data are consistent with the theory as to who was most affected by the stigma. I do this in three ways.

RQ2a. Are the relationships between CEP participation and student outcomes moderated by previous free-and-reduced-price-meal (FRPM) eligibility status – (a) students who were never eligible for FRPM, (b) students who were previously eligible for reduced-price meals, (c) students who were previously eligible for free meals, and (d) students who were not eligible for FRPM in the year prior to CEP but were in year(s) prior?

It is reasonable to expect that the results of research question 1 may vary according to students' prior FRPM eligibility circumstances. To address research question 2, I will examine these possible heterogeneous effects, disaggregating students into one of four groups: (a) students who pre-CEP were never eligible for FRPM, (b) students who pre-CEP were eligible for reduced prices for school meals, (c) students who pre-CEP were eligible for free meals, and (d) students who were not eligible for FRPM in the year prior to CEP but were previously eligible.

Students in group (a) were never eligible to receive free or reduced-price meals prior to implementation of CEP. However, under CEP, these students now have access to free meals. If this food is a better alternative than their previous meal options, I would hypothesize a positive impact on student outcomes. On the other hand, if the food is a worse alternative, I would expect a negative impact or no effect at all if students do not partake in school meals and continue with their previous meal option.

Students in group (b) were previously eligible to pay reduced prices for school meals. Under CEP, they are now eligible to eat for free. Prior research has shown that school meal price decreases typically result in increased school meal participation (Gleason, 1996; Gordon, Crepinsek & Condon, 2007; Maurer, 1984). Therefore, I hypothesize a positive impact on these students.

Students in group (c) can be split into two subgroups. While all students in this group have completed the required application and have always been eligible to receive free meals, one subgroup of these students act in accordance with their eligibility status and partake in free school meals, which I term "accepters"; the other subgroup of students are influenced by the social stigma of receiving FRPM and choose not to participate in receiving free meals, which I term as the "stigmatized" subgroup. Because individual meal participation data is not available, it is impossible to distinguish the two subgroups from one another. Therefore, to the extent that group (c) is comprised of "accepters," I hypothesize no impact of CEP on these students. To the extent that group (c) is comprised of the "stigmatized," I might expect a positive impact on student outcomes. Likely, these two groups will be mixed, in which case larger proportions of stigmatized students would increase the effects on student outcomes but larger proportions of "accepters" would attenuate the effects.

Group (d) consists of another subgroup of students who might feel stigmatized by participating in FRPM. The students in this group were previously eligible, but were not in the year prior to CEP. This may be due to one of two reasons. Some of these students chose not to complete the application because of the social stigma attached with it. Others may have experienced a change in family income that no longer qualified them for FRPM. Because family income data is unavailable, it is difficult to distinguish one subgroup from the other. Again, to the extent that group (d) is comprised of the "stigmatized" students, I hypothesize a positive impact on student outcomes. To the extent that the group is comprised of students who genuinely changed family income status and were no longer eligible, I might expect effects similar to that of group (a) – positive if the school meal option is a better alternative than their other options, negative if it is worse, or no effect if the students do not participate in free school meals.

RQ2b. Do the relationships between CEP participation and student outcomes vary by school level?

Previous research indicates that older students are more influenced by the social stigma of FRPM and less likely to participate in school meals than younger children (Dahl & Scholz, 2011). Therefore, I hypothesize that students in middle and high school have a greater proportion of stigmatized students than elementary school students. As a result, I should see a greater (or more positive) impact on middle and high school students.

RQ2c. Is the relationship between CEP participation and student outcomes moderated by the percentage of students in a school that were individually eligible for FRPM?

In order for the stigma to exist, there must be a group of students that are not eligible for free meals in the prior system. Students in schools with fewer students eligible for FRPM status would likely suffer more from the social stigma and benefit more from CEP than students in schools in which a greater proportion of students are eligible for free meals. This logic is substantiated in an interview from a prior study (Glantz et al., 1994), "It makes a difference in the area that you live in. In this area, I think about 80% get it, so it's not that big of a ridicule kind of thing for my kids." RQ2c seeks to identify if the results of research question 1 are also moderated by a school-level characteristic, the percentage of students that were previously eligible for FRPM. Based on this logic, fewer previously eligible students likely leads to greater concerns about stigma, and therefore, greater effects of CEP.

These three approaches indicate three different subpopulations likely affected by the stigma. To the extent that differential effects are found for each of the above subgroups, we can be more confident in CEP having successfully reduced the social stigma with participating in free school meals.

Methods

Empirical Framework

An ideal strategy for assessing the CEP program would be to randomly assign schools to the treatment and control groups. The difference in outcomes between students in the treatment, those attending schools participating in CEP, and students in the comparison group, those attending schools not participating in CEP, would yield the impact of the CEP program. However, this approach is not practical given the context. Alternatively, a regression discontinuity design can make use of the eligibility criteria to assess the impact of the CEP program at the threshold separating those eligible from those ineligible. However, several complications make the regression discontinuity approach problematic. For a discussion examining the problems of this approach, see the Appendix.

For this study, I use what I believe is the next best alternative, a comparative interrupted time series design. Comparative interrupted time series (CITS) designs have recently been used to evaluate education and employment programs (Bloom & Rico, 2002; Somers et al., 2013; St. Clair, Hallberg, and Cook, 2016). CITS compares the outcomes of the treatment and comparison groups pre- and post-intervention, which allows for an analysis of gains while controlling for secular trends. The procedure is similar to a difference-in-differences (DD) approach except that the pre-intervention time trends in the treatment and comparison groups are modeled. Therefore, effect estimates are determined by assessing whether the treatment group deviates from its baseline trend by a greater amount than the comparison group differs from its baseline trend after the treatment has begun. To model these trends, CITS designs have greater data requirements. At least four years of pre-treatment data are necessary (Somers et al., 2016).

Relative to a DD approach, an advantage of the CITS is that the key assumption behind the DD approach can be relaxed. In a DD, the parallel trends assumption posits that the average change in the treatment group would be the same as the comparison group if there were no treatment given covariates. A CITS directly accounts for pre-intervention trends, allowing the treatment and control groups' trends to vary in the period prior to treatment. Nonetheless, parallel trends can substantiate the CITS approach by providing confidence that the treatment and control groups react similarly to the treatment. While this assumption cannot be directly tested, I conduct validity checks to provide more confidence in the parallel trends assumption in two ways – graphically examining pre-CEP trends and fitting least square regression lines to the pre-CEP outcomes and comparing slopes. I also test for changes in student population as a function of the CEP program to assess whether the CEP program may have affected school populations. To the extent that the validity checks are passed, the CITS can provide plausibly causal estimates of program impacts.

Data & Measures

This study relies on two main data sources. The first is a publicly available dataset from the Tennessee Department of Education (TDOE) and provides the main independent variables for the analysis – the identified percentage of students (ISP) who qualify for FRPM under the CEP program for each district and each school and an indicator of whether these entities chose to participate in the CEP program. This data is available for the first three years of CEP – 2014-15 to 2016-17. While schools that were eligible in the first year did not have to reapply for CEP status until after four years, the data still provides this information for each school and each district in each year.

The second dataset is a rich student-level administrative dataset for Tennessee from the 2009-10 to 2016-17 school years. It is provided by a partnership between TDOE and the Tennessee Education Research Alliance (TERA) at Vanderbilt University. The outcome variables include student attendance, enrollment, on-time grade promotion, discipline, and achievement. Student enrollment will be calculated at the school level in the same way enrollment is calculated at TDOE – using average daily membership (ADM). For each day in the school year, the total number of enrolled students is tallied. The average is then calculated across the entire school year.

To evaluate the effects of CEP on student achievement, I use student standardized test scores. In Tennessee, all students in grades three to eight are tested on an annual basis in reading, math, and science using the Tennessee Comprehensive Assessment Program (TCAP). High school students take end of course (EOC) exams upon completion of select courses (English I, English II, English III, Algebra I, Geometry, Algebra II, Biology, and Chemistry) rather than a particular grade. Students receive scale scores for each of these exams. For this analysis, TCAP scale scores are standardized by year and grade; EOC scores are standardized by year and subject.

While student achievement measures are only available for students in tested grades, the effects of CEP on on-time grade progression can be broadened to include students in Kindergarten through 12th grade. On-time grade progression will be operationalized as a binary indicator of whether student i in grade g progressed to at least grade $g + 1$ in the following school year for grades Kindergarten through 11th grade. For 12th grade students, on-time grade progression will be operationalized as whether student *i* graduated high school that year.

The remaining dependent variables capture students in all grades. Student attendance will be operationalized as the percentage of days enrolled that a student attends school. Disciplinary offenses will be operationalized as a binary variable – whether a student committed an offense that resulted in a suspension or expulsion that year[2].

For research question 2, groups of students must be identified according to (a) students who pre-CEP were eligible for FRPM, (b) students who pre-CEP were eligible for reduced prices for school meals, (c) students who pre-CEP were eligible for free meals, and (d) students who were not eligible for FRPM in the year prior to CEP but were previously eligible, which I term "switchers". To assign students to each group, I start with FRPL statuses from the year prior to CEP start-up. I assume that the student would have continued in the same status in the first year of CEP had the program not gone into effect. Students who paid reduced prices for school meals in 2013-14 are coded as belonging to group (b). Students who received free school meals in 2013-14 are coded as belonging to group (c). If students did not receive free or reduced-price meals in 2013-14, they can fall into group (a) or (d). To distinguish between these two groups, I use data dating back to the 2009-10 school year. Students who never received free or reduced-price meals during their extent in the Tennessee public school system are coded as belonging to group (a). Students who received FRPM at some point prior to 2013-14 are coded as belonging to the "switchers" group (d)[3]. As previously discussed, this is not a perfect solution, as members

[2] Alternatively, the discipline offense outcome variable can be operationalized as the number of offenses per year. However, the results are difficult to interpret for this variable, i.e., a student attending a school participating in the CEP program is predicted to commit 0.04 less offenses, holding all else constant.
[3] Note that this status cannot be determined for all students – some students who entered a CEP school in the Tennessee public school system in 2014-15 or after do not consistently have a clear status beyond being in a CEP school. In cases where students do have a FRPM status, students entering a CEP school

of group (d) may have genuinely switched off FRPM status due to an increase in family income. However, this is the best alternative given lack of data on family incomes[4].

The TDOE/TERA dataset will also provide a number of other demographic characteristics to be used as control variables. This includes gender, race, special education status, English language learner status, and grade. A third dataset, the Common Core of Data, will provide school urbanicity codes and school level data (elementary, middle, high).

Sample

The analytic sample will include students in all treated non-alternative public schools in Tennessee. However, from the untreated schools, a sensible comparison group will need to be constructed. In an analysis comparing effect estimates from observational studies to estimates from random assignment, Bifulco (2012) shows that using comparison groups consisting of students drawn from districts with similar student body characteristics as the districts where treatment group students reside can reduce the bias in non-experimental methods by 64 to 96 percent. Therefore, I include in the comparison group non-alternative public schools in all districts that qualified for CEP as a whole but chose not to participate. Because many districts that did not qualify for CEP as a whole had multiple schools that did qualify, it does not make sense to only include other CEP-eligible districts. I also include in the comparison group all CEP-eligible schools who did not participate in the program. This broadens the group of comparison schools and ensures that the treatment and comparison groups include similar students with regards to school-level poverty rates.

---

may be less motivated to report correctly as the incentive to report is no longer a product of their individual economic status. Therefore, these students are omitted for this analysis.

[4] As a robustness check, I operationalize the definition of switch to only include students who switched status in the prior year only, in the prior two years, and in the prior three years. These various definitions yield consistent findings.

It is important to examine whether the comparison schools serve as a strong counterfactual for the treated CEP schools. In other words, would the comparison group respond to the CEP program in the same way the treatment group did had the comparison group implemented the CEP program? While this cannot definitely be proven, checking for balance across the two groups can provide confidence in the assumption. Table 1 shows baseline characteristics of schools in the treatment and schools in the described comparison group in 2013-14, the year prior to the first year of CEP. The two groups are similar in the percentage of male students and English language learners, and while statistically different, the difference in the percentage of special education and economically disadvantaged students is not practically large. The latter is encouraging given that the treatment favors schools that have large proportions of economically disadvantaged students. The major differences between the treatment and comparison groups lie in the percentage of students of color and the prior standardized test scores – students in the treatment group have a greater proportion of students of color and performed worse on standardized tests relative to the comparison group in the 2013-14 year. Additionally, the majority of schools are in the treatment group (as highlighted in figure 1), and urban and suburban schools are particularly overrepresented. Therefore, it will be important to control for these characteristics in the formal analyses.

Note that table 1 only compares the two groups on a small set of observable characteristics. The two groups may also differ on unobservable characteristics. For instance, the fact that those in the treatment chose to participate in the CEP program and many of those in the comparison group chose not to participate may threaten the parallel trends assumption and serves as a limitation to this study.

Table 1. Baseline Characteristics of Treatment and Control Groups

School Characteristic	Treatment Group	Comparison Group
Male	51%	51%
Students of Color	**40%**	**21%**
Economically Disadvantaged	**69%**	**65%**
Special Education	**15%**	**16%**
English Language Learners	4%	3%
Reading Standardized Score	**-0.16**	**-0.02**
Math Standardized Score	**-0.13**	**0.01**
Science Standardized Score	**-0.18**	**0.02**
Urban	36% (434 schools)	12% (28 schools)
Suburban	11% (136 schools)	6% (15 schools)
Town	13% (157 schools)	18% (42 schools)
Rural	40% (483 schools)	64% (153 schools)

Note: Statistically significant differences are indicated in bold. The number of schools of each location type is not tested.

Based on outcome measures, analytic samples will also vary. Attendance and discipline infractions are available for all students in all grades. On-time grade promotion data is unavailable for the last year of available data (2016-17) as it requires data from the following year. Standardized test scores are only available for students in grades three and up. Kindergarten through second grade students will be omitted from analyses assessing the effect of CEP on test scores. Third grade students will also be excluded as the model requires a lagged test score that is not available for students in their first year of testing. Lastly, test score analyses only include one year of test scores in 2014-15, the first year of CEP. In 2015-16, Tennessee experienced an anomalous year in which most students did not complete standardized testing. Therefore, student assessment data is not available for year two of CEP and pretest scores that are required for value-added estimates of the effects on student achievement are unavailable for year three.

Analytic Strategy

I model the comparative interrupted time series as follows:

$$y_{igst} = \beta_0 + \beta_1\, ever\_treat_s + \beta_2\, relyear_t + \beta_3\, ever\_treat_s \cdot relyear_t + \beta_4\, after_t + \beta_5\, ever\_treat_s \cdot after_t + X_{st}\mathrm{B_j} + S_{ist}\mathrm{B_k} + \gamma_g + \delta_t + \theta_s + e_{igst} \quad (1)$$

y_{igst} represents the dependent variable for student *i* in grade *g* school *s* in year *t* (the dependent variables include attendance, whether the student committed a discipline infraction that resulted in a suspension or expulsion, on-time grade progression, exit status, and standardized test scores). $ever\_treat_s$ is a binary variable indicating whether school *s* ever participated in CEP, regardless of year. $after_t$ is a binary variable indicating the years school *s* participated in CEP. For comparison schools and schools beginning CEP in the first year it is available, 2014-15, this variable takes a value of one for every year beginning in 2014-15 and zero for years prior to 2014-15. For schools that started CEP in 2015-16, this variable is adjusted so that it takes a value of zero in 2014-15. Likewise, for schools that started CEP in 2016-17, this variable is adjusted so that it takes a value of zero in 2014-15 and 2015-16. $relyear_t$ is a continuous variable for the year centered at the baseline year. For schools starting CEP in 2014-15, this variable is centered with $relyear_t$ equal to zero in 2013-14. For schools starting CEP in 2015-16, this variable is centered with $relyear_t$ equal to zero in 2014-15. And for schools starting CEP in 2016-17, this variable is centered with $relyear_t$ equal to zero in 2015-16. A value of one then represents the first year of CEP, two the second year of CEP, and so forth. A value of 0 represents the year prior to CEP, negative one two years prior to CEP, and so forth. The $relyear_t$ variable allows for time trends in the model, a key characteristic of the CITS design.

X_{st} is a vector of school level characteristics, including percent minority, percentage of economically disadvantaged students, percent English language learners, percent special education, school urbanicity, school level (elementary, middle, high), and whether the school is a charter school. S_{ist} serves as a vector of student characteristics for student *i* in school *s* in year *t* and includes gender, race, economically disadvantaged status, special education status, and English language learner status. In models assessing test scores as an outcome, I also include a lagged test score control variable as previous research finds that including pretreatment test scores can dramatically reduce bias in quasi-experimental methods (Bifulco, 2012; Shadish, Clark & Steiner, 2008). The inclusion of the lagged test score also allows for a value-added interpretation of the outcome. γ_g allows for a grade fixed effect, as students in various grades may respond differently to the CEP program, and because different schools participate in CEP in different years, δ_t serves as a year fixed effect to control for time-varying characteristics. θ_s serves as a school fixed effect, which accounts for differences in how schools may react to the CEP program.

Based on these variables, β_0 represents the baseline mean (intercept) for the comparison schools. $\beta_0 + \beta_1$ represents the baseline mean (intercept) for CEP-participating schools. β_2 represents the baseline slope for the comparison schools. $\beta_2 + \beta_3$ represents the baseline slope for CEP-participating schools. β_4 represents the deviation from the baseline trend for the comparison schools in the years of CEP. $\beta_4 + \beta_5$ represents the deviation from the baseline trend for CEP-participating schools in the years participating in CEP. The key coefficients of interest are $\beta_5 + \beta_3$ * *year*, where *year* indicates the year of treatment (in this case, one, two, or three). If β_3 is not significant, the average treatment effect of CEP is represented by β_5 itself.

It is possible that the effect of the CEP program varies by year of implementation. For instance, schools may have adjusted the meal distribution process as they better gauged the number of students choosing to participate in the free meal program. Therefore, I expand equation (1) to assess the yearly impact of the CEP program in the following form:

$$y_{ist} = \beta_0 + \beta_1\, ever\_treat_s + \beta_2\, relyear_t + \beta_3\, ever\_treat_s \cdot relyear_t + \beta_4\, year1_t + \beta_5\, ever\_treat_s \cdot year1_t + \beta_6\, year2_t + \beta_7\, ever\_treat_s \cdot year2_t + \beta_8\, year3_t + \beta_9\, ever\_treat_s \cdot year3_t + \beta_{10}\, y_{ist-1} + X_{st}\mathrm{B}_j + S_{ist}\mathrm{B}_k + e_{ist} \quad (2)$$

In this case, β_5, β_7 and β_9 represent the average treatment effect of those under CEP in their first, second, and third years, respectively.

Lastly, in conjunction with research question 2, I will assess for moderating effects using samples limited to the respective sub-populations.

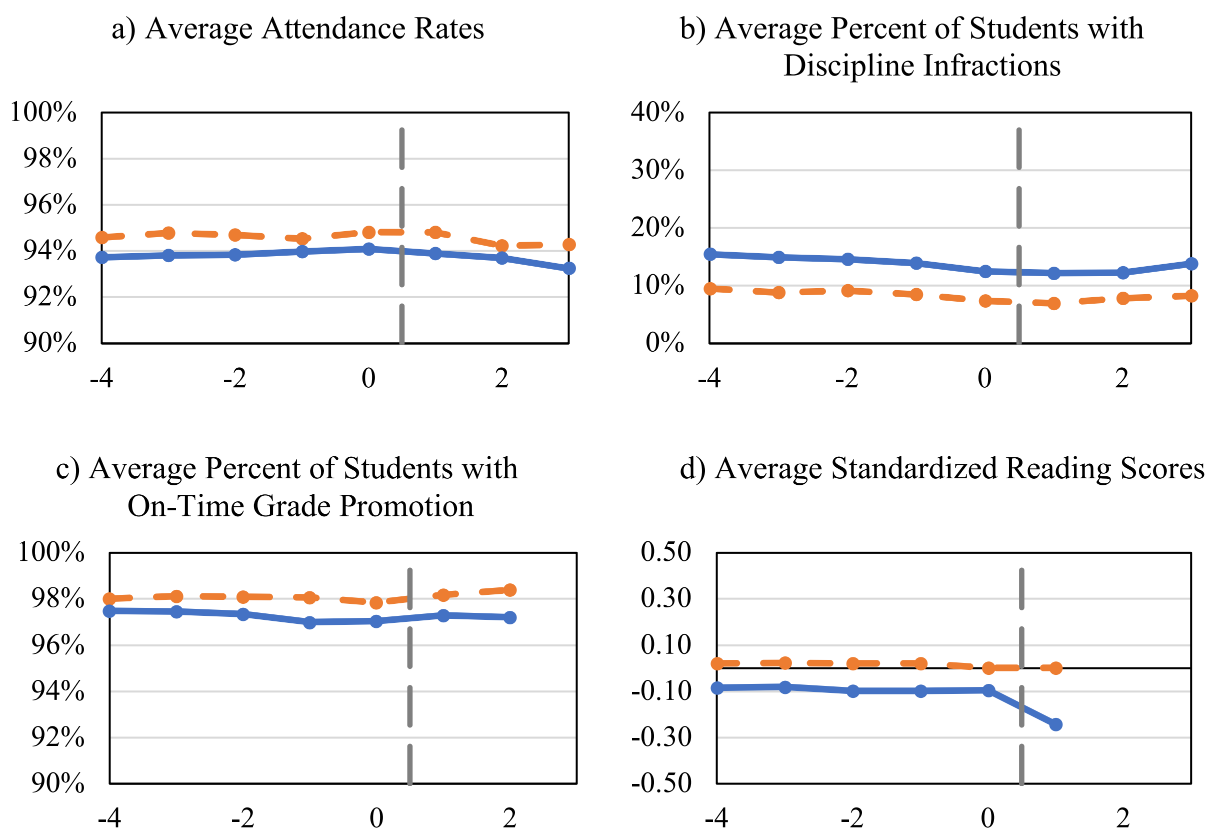
Results

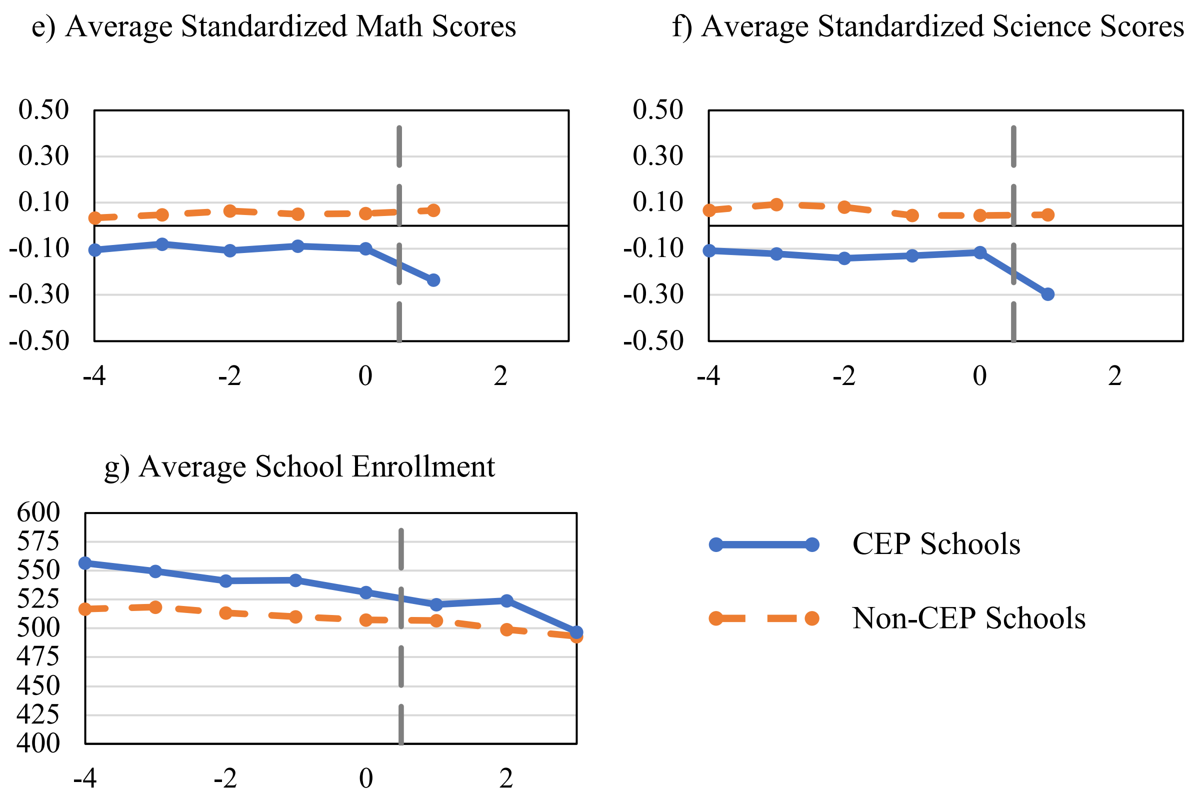
In this section, I present the results of the evaluation of the Community Eligibility Provision in Tennessee. As previously discussed, the parallel trends assumption can be relaxed with a CITS design. However, parallel trends can provide greater confidence that the treatment and comparison groups respond similarly to the treatment. Therefore, I first check the validity of the parallel trends assumption. Following this, in conjunction with research question 1, I present the overall effect of the Community Eligibility Provision, both over the course of all three years and year-by-year. I then answer research question 2 on the effects of CEP on students who are most likely to be affected by the social stigma of participating in the school free meals program by examining the effects by students' prior FRPM status, school level, and school FRPM percentage.

Checking the Parallel Trends Assumption

To assess the parallel trends assumption, I first examine pre-intervention trends of the two groups. Outcome trends that are parallel prior to treatment provides support that the trend would be similar after treatment, except for the effect of the CEP program. Figure 2 graphically depicts the average trends for each outcome across the five years prior to the intervention. With the exception of school enrollment, the trends for both CEP and non-CEP schools appear to be fairly flat across all outcomes. For average school enrollment, the slope of the pre-intervention trend appears to be slightly more negative for CEP schools than for non-CEP schools. Overall, the pretreatment trends provide support for the parallel trends assumption with some slight concerns for average school enrollment. Additionally, the fact that the trends tend to be fairly linear across the board also provides greater support for the CITS approach.

Figure 2. Pre-Treatment Trends





In a second check for parallel trends, I fit a least squares regression line to each of the pre-intervention trends displayed in Figure 2. The slopes of those lines are displayed in table 2 separately for the treatment and comparison groups. The p-value for a chi-squared test of significant differences across the two slopes is also provided. The slopes across both CEP and non-CEP groups are nearly zero in all cases, except for school enrollment. Consequently, the difference in slopes are also nearly zero. Significant differences in slopes are found in the outcomes of offender, math score, and school enrollment. These small but significant differences provide further support for the CITS approach, which models pre-intervention trends for CEP and non-CEP schools.

Table 2. Slopes of Pre-Intervention Trends

Outcome	Slope of Non-CEP Schools	Slope of CEP Schools	Difference in Slopes	p-value
Attendance Rate	0.0003	0.0007	0.0004	0.218
Offender	-0.0029	-0.0050	-0.0021	0.041
On-time Grade Promotion	-0.0008	-0.0013	-0.0005	0.471
Reading	-0.0035	-0.0078	-0.0044	0.290
Math	0.0065	-0.0079	-0.0143	0.010
Science	-0.0082	-0.0112	-0.0030	0.468
School Enrollment	-0.0060	-0.0506	-0.0446	0.012

Main Effects

I first descriptively examine whether the CEP program was effective in expanding access to free school meals to more students. Figure 3 displays the percentage of all Tennessee students who qualified for, or had access to, free meals (dark gray) and the percentage of students who qualified for reduced-price meals (light gray) each year included in the analysis. (This analysis assumes that schools participating in CEP provided free school meals to all students as specified by the program.) Before the implementation of CEP, indicated by the vertical dashed line, the rates for access to free meals were slowly and steadily increasing from about 47% in 2009-10 to 54% in 2013-14. During this time period, the percentage of students with access to reduced price meals remained fairly stable at about 5-6%. In the first year CEP became available, access to free meals jumped approximately 12 percentage points to 66% of students in Tennessee. The percent of students accessing reduced price meals dropped about 2 percentage points, as many of these students now had access to free meals. In the following year, 2015-16, as more districts and schools adopted, the percentage of students with access to free meals increased another 14 percentage points to 80% of students in Tennessee. Two percent of students still had access to reduced price meals. These numbers remained stable through 2016-17 as there was little change in the number of districts or schools that were eligible and opted into the CEP program, as shown in Figure 1. With increased access to free school meals as a result of the CEP program, I assume greater student participation, which is supported by previous literature (Logan et al., 2014; Henry, 2015; Hong, 2015), and examine next whether this access translated to improved student educational outcomes.

Figure 3. Student Access to Free and Reduced Price Meals Over Time

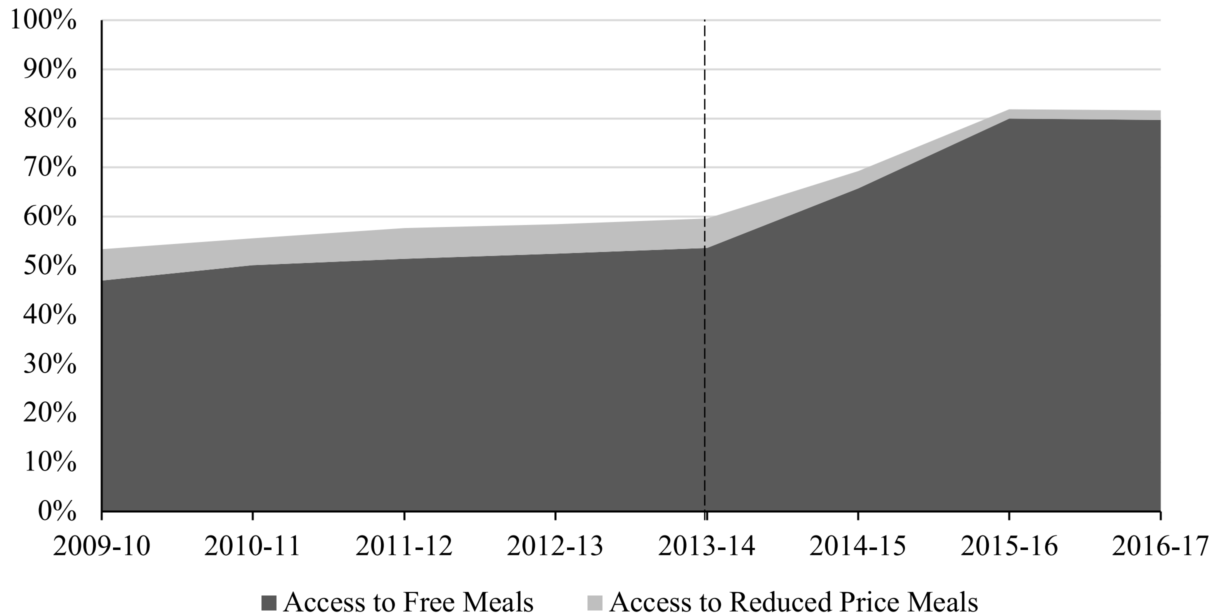


Table 4 provides estimates of the CEP program's effects on student attendance rates, whether students commit a disciplinable offense, whether a student is promoted to the next grade on time, reading, math, and science standardized test scores, and school-level enrollment.[5] In column 1, I find that students enrolled in CEP schools are, overall, 1.2 percentage points less likely to commit a suspendable/expellable offense than students in non-CEP schools. I do not find any effects on the other five student outcomes overall.

[5] The coefficient β_3 (in equation 1) is found to be statistically insignificant with all outcomes. Therefore, I only report the β_5 coefficient estimates for the overall effect here. Detailed results are available upon request.

As with many school reforms, the CEP program may have required multiple years to fully and correctly implement. Additionally, it may take time for the impact to be fully realized. To assess these possibilities, I conduct a year-by-year analysis of the CEP program, the results of which are displayed in columns 2 to 4 of table 4. As a reminder, standardized test scores were unavailable in 2015-16, which precludes Year 2 and 3 analyses for assessments, and third year on-time grade promotion data are unavailable as these require year 4 data to assess. Though statistically insignificant overall, in the third year, I find a small negative effect on attendance rates that translates to about half of a school day.

The effect of CEP on the likelihood of students committing disciplinable offenses increases over time, with students becoming less likely to commit offenses as CEP implementation continues. In year one, the effects are null but grow to 1.5 percentage points less likely to commit an offense in year 2 and to 2.3 percentage points less likely to commit an offense in year 3.

A similar pattern is exhibited with on-time grade promotion. In year 1, the effect of CEP on on-time grade promotion is null. However, by year 2, students in CEP schools are 0.6 percentage points more likely to be promoted to the next grade on time. It is possible that this positive trend continues into year 3 as it does with committing a disciplinable offense, but that is not assessable with this data. Unfortunately, test score results are unavailable for years 2 and 3 – therefore I am unable to assess whether the effects of CEP on student achievement is realized over a longer period after the program was implemented.

Table 3. Effects of the CEP Program

Outcome	Overall	Year by Year Analysis		
		Year 1	Year 2	Year 3
Attendance Rate	-0.000	-0.001	-0.000	-0.003*
	(0.001)	(0.001)	(0.001)	(0.002)
N	6,562,896			
Offender	-0.012*	-0.007	-0.015**	-0.023**
	(0.005)	(0.005)	(0.006)	(0.008)
N	6,578,479			
On-Time Grade Promotion	0.002	0.001	0.006*	-
	(0.002)	(0.002)	(0.003)	
N	5,460,842			
Reading Score	-0.003	-	-	-
	(0.009)			
N	2,085,842			
Math Score	-0.007	-	-	-
	(0.016)			
N	1,811,542			
Science Score	0.013	-	-	-
	(0.014)			
N	1,723,351			
School Enrollment (in 100's)	0.022	0.018	0.008	0.015
	(0.043)	(0.042)	(0.052)	(0.067)
N	11,265			

Standard errors are shown in parentheses. Standard errors are clustered at the school level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All models include the following school-level covariates - school percent minority, school percent economically disadvantaged, school percent English language learners, school percent special education, school level, charter indicator, urbanicity. All models except those evaluating the impact of CEP on school enrollment also include student-level covariates - male, race, economically disadvantaged status, English language learner status, and special education status. All models also include grade fixed effects, year fixed effects, and school fixed effects, with the exception that models evaluating the impact of CEP on school enrollment do not include grade or school fixed effects.

Note that the results are fairly consistent across the overall and year-by-year models. Therefore, for brevity's sake, I focus on the overall models moving forward.

Though the outcomes appear fairly linear in figure 2, incorrectly modeled trends can bias the effect estimates of a CITS (Candelaria & Shores, 2018). Therefore, I conduct a robustness check of the overall results above using a different ITS approach, a difference in differences (DD) model, which removes all formal modeling of trends. For sake of space, I include the results in table 1 of the appendix. I find consistent substantive conclusions between the CITS and DD models.

Assessing the Effects on Stigma

To explore the impact of CEP on stigma, I first examine heterogeneous effects based on students' prior FRPM status. I assign students into one of four categories based on their status in the year prior to CEP implementation and years previous – never free or reduced-price meals, reduced price meals, free meals, and switchers (students who were not on free or reduced price meals in the year prior to CEP implementation, but had at some point in the previous four years received FRPM) – and assess the effects separately across each group of students. The results are displayed in Table 5. Students who were never eligible for free or reduced-price meals were 0.8 percentage points less likely to commit an offense as a result of the CEP program. However, there were no effects on attendance, on-time grade promotion, or test scores for these students. Students who previously received free meals experienced a 0.05 percentage point increase in the likelihood of progressing onto the next grade on time. However, they were also less likely to attend school 0.2% of the school year, which translates to less than half of a school day. No other effects were found for these students who previously had access to free meals. No effects were found for students who were previously on reduced-price meals or for students who switched FRPM status.

Table 4. Effects of the CEP Program by Student Prior Free & Reduced-Price Meal Status

Outcome	Student FRPM Status			
	Never FRPM	**Reduced-Price Meals**	**Free Meals**	**Switchers**
Attendance Rate	0.000	-0.001	-0.002*	-0.002
	(0.001)	(0.001)	(0.001)	(0.003)
N	1,227,540	332,045	3,041,548	501,113
Offender	-0.008*	-0.003	-0.003	-0.010
	(0.003)	(0.005)	(0.005)	(0.008)
N	1,231,037	332,750	3,049,085	502,390
On-Time Grade Promotion	0.000	-0.000	0.005**	0.006
	(0.002)	(0.004)	(0.002)	(0.004)
N	1,060,880	289,958	2,624,428	443,914
Reading Score	-0.010	-0.013	-0.018	-0.017
	(0.012)	(0.016)	(0.010)	(0.015)
N	454,751	126,772	943,297	216,006
Math Score	-0.012	-0.021	-0.020	-0.029
	(0.020)	(0.022)	(0.017)	(0.023)
N	409,874	115,533	865,324	194,179
Science Score	0.022	0.025	-0.006	-0.006
	(0.017)	(0.021)	(0.015)	(0.021)
N	395,927	111,341	831,362	186,108

Standard errors are shown in parentheses. Standard errors are clustered at the school level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All models include the following school-level covariates - school percent minority, school percent economically disadvantaged, school percent English language learners, school percent special education, school level, charter indicator, urbanicity. All models also include student-level covariates - male, race, economically disadvantaged status, English language learner status, and special education status. All models also include grade fixed effects, year fixed effects, and school fixed effects.

Next, I examine the effects of CEP by school level. The results are presented in table 6. No effects are found in elementary or middle schools. However, there are effects in high schools. It appears that the overall effect on discipline found in table 4 is driven by high schools. Students in high schools with CEP are 4.6 percentage points less likely to commit a disciplinable offense. No other effects are found in high schools.

Table 5. Effects of the CEP Program by School Level

Outcome	School Level		
	Elementary	**Middle**	**High**
Attendance Rate	-0.001	-0.001	0.002
	(0.001)	(0.002)	(0.004)
N	3,183,761	1,561,807	1,726,929
Offender	-0.001	-0.006	-0.046**
	(0.003)	(0.009)	(0.015)
N	3,192,468	1,565,408	1,730,173
On-Time Grade Promotion	0.001	0.004	0.000
	(0.001)	(0.002)	(0.009)
N	2,670,581	1,306,570	1,411,485
Reading Score	-0.018	0.000	0.021
	(0.015)	(0.014)	(0.016)
N	575,469	883,484	586,309
Math Score	-0.019	0.008	0.004
	(0.022)	(0.024)	(0.041)
N	576,103	883,762	330,227
Science Score	0.024	0.013	0.010
	(0.020)	(0.021)	(0.032)
N	575,141	881,689	248,700
School Enrollment (in 100's)	-0.000	0.130	-0.199
	(0.049)	(0.074)	(0.142)
N	6,412	2,743	1,934

Standard errors are shown in parentheses. Standard errors are clustered at the school level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All models include the following school-level covariates - school percent minority, school percent economically disadvantaged, school percent English language learners, school percent special education, school level, charter indicator, urbanicity. All models except those evaluating the impact of CEP on school enrollment also include student-level covariates - male, race, economically disadvantaged status, English language learner status, and special education status. All models also include grade fixed effects, year fixed effects, and school fixed effects, with the exception that models evaluating the impact of CEP on school enrollment do not include grade or school fixed effects.

Lastly, I examine whether the relationships between CEP status and student outcomes are moderated by the percentage of students in a school that were individually eligible for FRPM, or school economically disadvantaged status. The results of this analysis are displayed in table 7. Three of the outcomes yield statistically significant moderating effects. For every ten percentage point increase of students eligible for FRPM at a school, CEP reduces attendance rates by 0.24%. In other words, for every ten percentage point decrease of economically disadvantaged students, CEP increases attendance rates by 0.24%, which translates to approximately half of a school day. Similarly, for every ten percentage point decrease in economically disadvantaged students, students in CEP schools are 0.88 percentage points less likely to commit a disciplinable offense and are predicted to score 0.028 standard deviations greater on the standardized science assessment. No moderating effects (by school economically disadvantaged percentage) were found for on-time grade promotion, reading scores, math scores, or total school enrollment.

Table 6. Assessing Moderating Effects of School ED Percentage on the CEP Program

Outcome	Effect Estimate	Standard Error	N
Attendance Rate	-0.024**	(0.008)	6,562,896
Offender	0.088**	(0.032)	6,578,479
On-Time Grade Promotion	0.003	(0.008)	5,460,842
Reading Score	-0.122	(0.084)	2,085,842
Math Score	0.002	(0.150)	1,811,542
Science Score	-0.276*	(0.123)	1,723,351
School Enrollment (in 100's)	0.253	(0.288)	11,265

Standard errors are shown in parentheses. Standard errors are clustered at the school level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. All models include the following school-level covariates - school percent minority, school percent economically disadvantaged, school percent English language learners, school percent special education, school level, charter indicator, urbanicity. All models except those evaluating the impact of CEP on school enrollment also include student-level covariates - male, race, economically disadvantaged status, English language learner status, and special education status. All models also include grade fixed effects, year fixed effects, and school fixed effects, with the exception that models evaluating the impact of CEP on school enrollment do not include grade or school fixed effects.

Discussion

In 2014-15, Tennessee changed the structure of its breakfast and lunch programs to allow schools with larger percentages of economically disadvantaged students to provide free meals to all students, regardless of income level. One of the large motivators for this change was to reduce stigma attached with participating in the free school meals program. This program increased access to free school meals by 50%, which in itself is already a positive outcome for ensuring all students have the ability to eat during the school day. This study examines how adoption of the Community Eligibility Provision impacted other student educational outcomes, including attendance, discipline, on-time grade progression, test scores, and school-level enrollment. Previous research on universal free meal programs have found that participation in the meal programs have increased (Leos-Urbel et al., 2013; Murphy et al., 1998; Bernstein et al., 2004; Schanzenbach & Zaki, 2014), but have yielded mixed results on student outcomes (Wahlstrom & Begalle, 1999; Kleinman et al., 2002; Lent & Emerson, 2007; Murphy et al., 1998; Bernstein et al., 2004; Schanzenbach & Zaki, 2014; Leos-Urbel et al., 2013; Ribar & Haldeman, 2013). The results of this study suggest that, overall, the Community Eligibility Provision decreased the likelihood of students committing a disciplinable offense by 1.2 percentage points overall with the greatest impact on high school students – by 4.6 percentage points. In addition, a year-by-year analysis showed that the effects of CEP on student educational outcomes grew over time, with more positive effects on on-time grade progression and behavioral outcomes in future years of implementation. In particular, students in CEP schools were 0.6 percentage points more likely to be promoted to the next grade on time in year two, 1.5 percentage points less likely to commit an offense in year two, and 2.3 percentage points less likely to commit an offense in year three. Unfortunately, year two and three data were unavailable for standardized test scores to assess whether similar patterns may have held.

By examining the effects of CEP on various other subgroups of students, this study also provides some suggestive evidence about whether the CEP program was able to reduce stigma of participation in school meal programs. In addition to the positive effects on discipline in high school students, a subgroup that has been found to be affected greater by the stigma, the program was found to increase the likelihood of on-time grade progression of students who were pre-CEP eligible for free meals. Within this group includes both students who formerly did not participate in the school meal program due to the stigma and students who did not suffer from the stigma and did participate. Without knowing the composition of this group, it is not clear whether CEP helped to reduce stigma. However, the CEP program, theoretically, should not have had an effect on non-stigmatized students since nothing changed for these students.

For stigma to be a concern, there must be a proportion of students who do not qualify for free meals. In particular, as the proportion of economically disadvantaged students decreases, students eligible for FRPM are more likely to experience the stigma as fewer students schoolwide are eligible to participate. If all students are eligible, it is less likely that students are stigmatized for participating in free meals. Therefore, outcomes that become more positive as the school percentage of economically disadvantaged students decreases indicate evidence consistent with CEP reducing stigma associated with free school meal participation. In an analysis examining the effect of CEP based on school economically disadvantaged percentages, I find that CEP increased attendance rates and science scores and decreased disciplinable offenses as the school-level percentage of economically disadvantaged students decreased. This collection of results suggests that CEP did help to reduce the stigma associated with participation in free school meals, leading to positive outcomes for these students.

Two results of this study suggests that the CEP program negatively impacted attendance rates. Students overall experienced a decrease in attendance rates in year three, as did the subgroup of students who pre-CEP were eligible for free school meals. However, in both cases, this effect was substantially small, translating to approximately half of a school day.

A few limitations should be considered in the light of this study. First, test score data for years two and three of CEP implementation could not be used because of testing complications.

Second, the structure of the program only permitted districts and schools with large enough proportions of ED students to participate in CEP. In addition, districts and schools that qualified but did not want to participate could choose not to, and schools that did not qualify but wished to participate could do so by partnering with other schools such that the group qualified. The ability to select in or out of the program provides some concerns between the equivalency of the treatment and comparison groups. However, the results of the assessment of parallel trends provides some support for comparing the two groups, and the use of a CITS approach allows for the two groups to differ in pre-intervention trends. Further research should examine why schools, particularly those who did not comply with eligibility status, chose to or not to participate in CEP.

Second, without knowing exactly the reasons that students chose to or not to participate in free school meals, I cannot definitively answer whether CEP reduced the associated stigma. However, the various approaches taken above considering student prior eligibility status, school level, and school level student economically disadvantaged rates provide evidence consistent with having reduced the stigma of participation.

Lastly, I lack the data to assess the extent to which CEP schools complied with the program and the quality in which the program was implemented. Without meal take-up rates, it is impossible to determine whether students actually participated in free school meals. In addition, measures of the quality of the food prepared are unavailable. The results of this study could be influenced by any changes in the nutritional value, the appeal, or the variety of options of the prepared meals. Further research should examine the uptake and the quality and quantity factors associated with meal options.

Overall, the results of this study suggest that the Community Eligibility Provision provided greater access to free meals during the school day, which extended to positive effects on behavioral outcomes particularly for high school students. It also appears that CEP may have also helped to reduce the stigma with participating in the free school meals program, at least in Tennessee. As of the 2016-17 school year, all states have provided the CEP option to their schools, but participation rates have ranged greatly with some states near 100% participation and other states as low as 20% participation (of eligible schools). Tennessee ranks 11th in the percent of eligible schools adopting CEP in the 2016-17 school year (FRAC, 2017). Given this, further research should be conducted on other states to assess the generalizability of this study's results. Collectively, these studies can help inform states with lower participation rates regarding the greater impact of the Community Eligibility Provision.

Appendix

A regression discontinuity approach was first proposed for this study. However, a number of complications arose making the approach problematic. First, there are multiple levels at which schools can qualify for CEP including the district level and the school level. This means that there are two different eligibility rules and therefore two different forcing variables that must be accounted for in the regression discontinuity model. Several studies (Wong, Steiner & Cook, 2013; Reardon & Robinson, 2012; Papay, Murnane & Willett, 2014; Porter et al., 2017; Cheng, 2016; Dee, 2012; Henry & Guthrie, 2015) have utilized a frontier regression discontinuity approach to address multiple forcing variables by estimating the impact at individual frontiers that restrict to conditions based on all but one forcing variable and evaluating these effects across multiple permutations of those restrictions. For instance, for the CEP program, one could restrict the sample to district-ineligible schools and estimate the impact on schools just eligible individually[6]. One disadvantage of this approach is that it further limits the statistical power to determine that an effect has occurred, which is already lower for regression discontinuity designs, and the extent to which the estimates of a regression discontinuity design can be generalized.

While this first concern is solvable, a second concern is more difficult to overcome. In many cases, schools and districts that were eligible for CEP chose not to opt in. This is entirely rational as at low eligibility rates, schools may have suffered a financial loss by joining the CEP program. Further, schools that were ineligible (or were eligible districtwide but the district chose not to participate) were allowed to formed groups of schools that could together apply for the

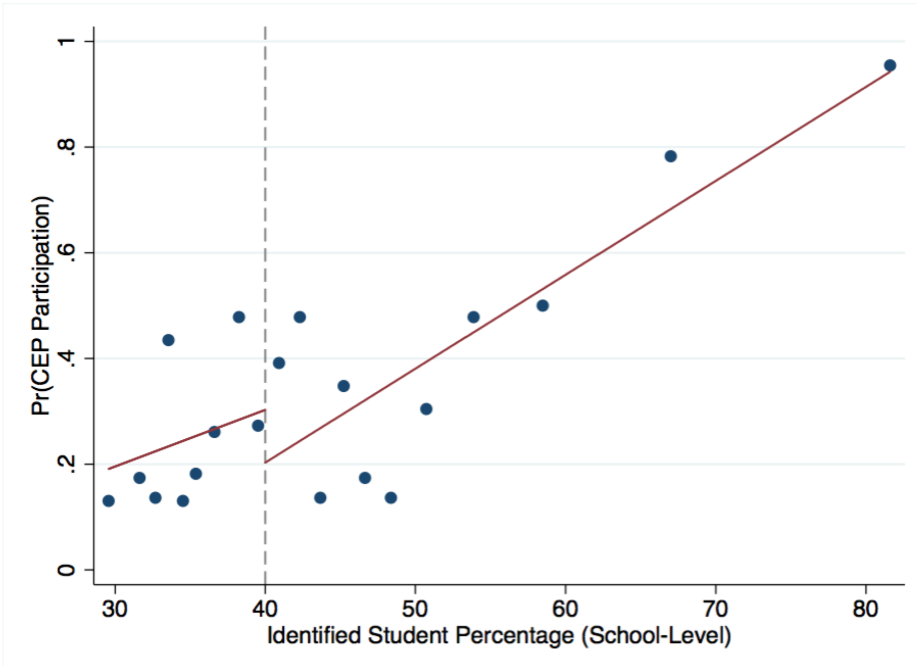
[6] The alternative – restricting the sample to school-ineligible schools and estimating the impact on districts just eligible – is infeasible. By removing school-ineligible schools from this analysis, this, in turn, changes the value of the district eligibility forcing variable, which is aggregated from the schools within the district.

CEP program, with little limitations on which schools could form groups. These two issues create concerns of noncompliance with the school eligibility regression discontinuity approach proposed earlier. Typically, such a concern could be addressed using a fuzzy regression discontinuity approach instrumenting participation in the program with eligibility and only estimating the effect on schools complying with eligibility status. However, compliance in eligibility for the CEP program, particularly near the threshold, is nearly indistinguishable in the eligible and ineligible schools. Figure 2 shows the probability of participation in CEP as a function of school level identified student percentages for schools in districts that are not eligible for districtwide participation for the first year of CEP. Figure 2a includes all schools across the full range of ISPs; figure 2b only includes schools in a smaller bandwidth – within 5 percentage points of the threshold. When using the full range of ISPs (in figure 2a), it appears that at the threshold, schools that are directly above and are eligible are less likely to participate in CEP than those directly below and ineligible. Using the smaller bandwidth (in figure 2b), schools that are directly above and are eligible are only 12-13 percentage points more likely to participate in CEP than those directly below and ineligible. To use the two stage least squares approach, the probability of take up has to be markedly higher in the eligible schools, which is not the case, as shown in the figures below. Additionally, for the restricted bandwidth, schools that are eligible are less likely to participate in CEP as their ISP increases, which runs contrary to the expected relationship between eligibility and participation. These figures illustrate that the noncompliance behind school and district eligibility rules prohibit a meaningful examination of the impact of the CEP program using a regression discontinuity approach at the eligibility threshold.

A third concern only extends the threat of noncompliance in years two and three of the CEP program. When schools or districts qualified for CEP and opted into the program, they remained eligible for the following four years. Thus, if the identified student percentage dropped below 40% for years 2, 3, or 4, the school or district could still participate if they had in year 1. This creates even greater noncompliance in years 2 and 3 of CEP, as many ISPs dropped below the threshold but these schools and districts continued to receive the treatment. This may also suggest possible manipulation of the data to ensure schools and districts met the criteria for participating in the program in at least one year.

Figure A1. Predicted Probabilities of District-Ineligible Schools Participating in CEP Based on Identified Student Percentages (ISPs)

(a) Including all district-ineligible schools



(b) Including those with ISPs within a 5 percentage point bandwidth of the threshold

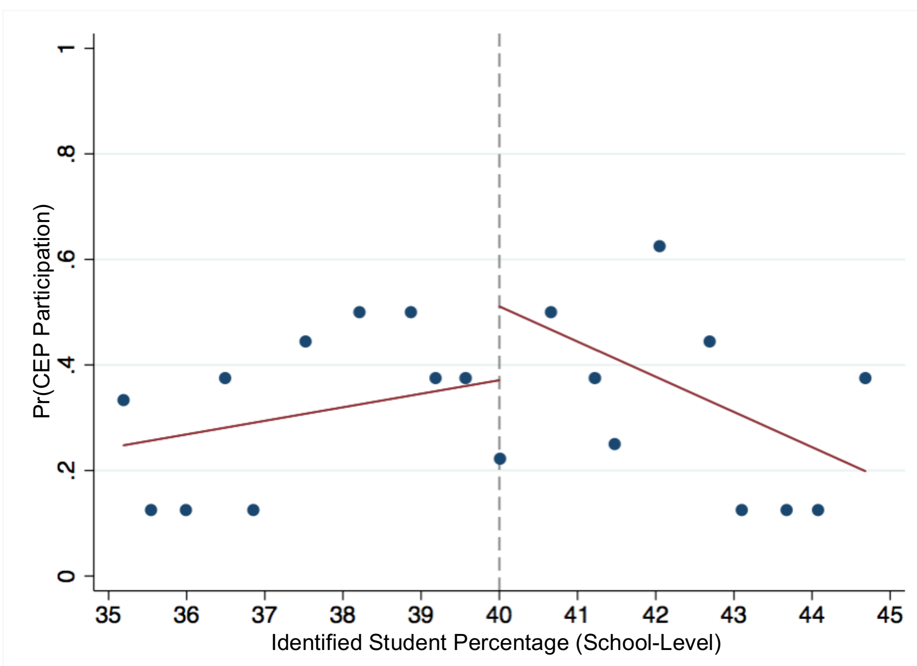


Table A1. Comparing CITS and DD Models Assessing the Effects of the CEP Program

Outcome	DD	CITS
Attendance Rate	-0.000	0.001
	(0.001)	(0.001)
N	6,562,896	
Offender	-0.012*	-0.009*
	(0.005)	(0.004)
N	6,578,479	
On-Time Grade Promotion	0.002	0.001
	(0.002)	(0.001)
N	5,460,842	
Reading Score	-0.003	0.004
	(0.009)	(0.007)
N	2,085,842	
Math Score	-0.007	-0.001
	(0.016)	(0.014)
N	1,811,542	
Science Score	0.013	0.014
	(0.014)	(0.013)
N	1,723,351	
School Enrollment (in 100's)	0.022	0.059
	(0.053)	(0.042)
N	11,265	

Standard errors are shown in parentheses. Standard errors are clustered at the school level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. DD = Difference in Differences, CITS = Comparative Interrupted Time Series. All models include the following school-level covariates - school percent minority, school percent economically disadvantaged, school percent English language learners, school percent special education, school level, charter indicator, urbanicity. All models except those evaluating the impact of CEP on school enrollment also include student-level covariates - male, race, economically disadvantaged status, English language learner status, and special education status. All models also include grade fixed effects, year fixed effects, and school fixed effects, with the exception that models evaluating the impact of CEP on school enrollment do not include grade or school fixed effects

References

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

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

Askelson, N. M., Golembiewski, E. H., DePriest, A. M., & Williams, S. C. (2015). Parent school breakfast survey: Report of findings.

Askelson, N. M., Golembiewski, E. H., Ghattas, A., Williams, S., Delger, P. J., & Scheidel, C. A. (2017). Exploring the Parents' Attitudes and Perceptions About School Breakfast to Understand Why Participation Is Low in a Rural Midwest State. Journal of nutrition education and behavior, 49(2), 107-116.

Bailey-Davis, L., Virus, A., McCoy, T. A., Wojtanowski, A., Vander Veur, S. S., & Foster, G. D. (2013). Middle school student and parent perceptions of government-sponsored free school breakfast and consumption: A qualitative inquiry in an urban setting. Journal of the Academy of Nutrition and Dietetics, 113(2), 251-257.

Bartfeld, J., & Kim, M. (2010). Participation in the School Breakfast Program: New evidence from the ECLS-K. Social Service Review, 84(4), 541-562.

Belot, M., & James, J. (2011). Healthy school meals and educational outcomes. *Journal of health economics*, *30*(3), 489-504.

Benton, D., & Parker, P. Y. (1998). Breakfast, blood glucose, and cognition. The American journal of clinical nutrition, 67(4), 772S-778S.

Bergman, E. A., Buergel, N. S., Enamuthu, J., & Sanchez, A. (2000). Time spent by schoolchildren to eat lunch. Journal of the Academy of Nutrition and Dietetics, 100(6), 696.

Bernstein, L. S., McLaughlin, J. E., Crepinsek, M. K., & Daft, L. M. (2004). Evaluation of the School Breakfast Program Pilot Project: Final Report. Special Nutrition Programs. Report Number CN-04-SBP. Nutrition Assistance Program Report Series. US Department of Agriculture.

Bhatia, R., Jones, P., & Reicker, Z. (2011). Competitive foods, discrimination, and participation in the National School Lunch Program. American journal of public health, 101(8), 1380-1386.

Bhattacharya, J., Currie, J., & Haider, S. J. (2006). Breakfast of champions? The School Breakfast Program and the nutrition of children and families. *Journal of Human Resources*, *41*(3), 445-466.

Bifulco, R. (2012). Can Nonexperimental Estimates Replicate Estimates Based on Random Assignment in Evaluations of School Choice? A Within-Study Comparison. *Journal of Policy Analysis and Management*, *31*(3), 729-751.

Bloom, H. S., & Rico, J. A. (2005). Using place-based random assignment and comparative interrupted time-series analysis to evaluate the jobs-plus employment program for public housing residents. *The ANNALS of the American Academy of Political and Social Science*, *599*(1), 19-51.

Candelaria, C. A., & Shores, K. A. (2017). Court-Ordered Finance Reforms in The Adequacy Era: Heterogeneous Causal Effects and Sensitivity. *Education Finance and Policy*.

Cheng, Y. A. (2016). Regression Discontinuity Designs with Multiple Assignment Variables. Retrieved from https://www.econ.berkeley.edu/sites/default/files/EconHonorsThesis_YizhuangAldenCheng.pdf.

Cohen, J. F., Jahn, J. L., Richardson, S., Cluggish, S. A., Parker, E., & Rimm, E. B. (2016). Amount of time to eat lunch is associated with children's selection and consumption of school meal entrée, fruits, vegetables, and milk. Journal of the Academy of Nutrition and Dietetics, 116(1), 123-128.

Cook, T. (2008). Waiting for life to arrive: A history of the regression-discontinuity design in psychology, statistics, and economics. Journal of Econometrics, 142, 636–654.

Dahl, M. W., & Scholz, J. K. (2011). The National School Lunch Program and school breakfast program: Evidence on participation and noncompliance. Madison, WI: Institute for Research on Poverty, University of Wisconsin-Madison. Retrieved from http://www.econ.wisc.edu/~ scholz/Research/Lunch.pdf.

Dee, T. (2012). *School turnarounds: Evidence from the 2009 stimulus* (No. w17990). National Bureau of Economic Research. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.716.7392&rep=rep1&type=pdf

Deke, J., & Dragoset, L. (2012). Statistical Power for Regression Discontinuity Designs in Education: Empirical Estimates of Design Effects Relative to Randomized Controlled Trials. Working Paper. *Mathematica Policy Research, Inc.*

Dunifon, R., & Kowaleski-Jones, L. (2003). The influences of participation in the National School Lunch Program and food insecurity on child well-being. Social Service Review, 77(1), 72-92.

Epple, D., & Romano, R. (2011). Peer effects in education: A survey of the theory and evidence. *Handbook of social economics*, *1*(11), 1053-1163.

Food Research & Action Center [FRAC]. (2017). Community Eligibility Continues to Grow in the 2016-2017 School Year. Retrieved from http://frac.org/wp-content/uploads/CEP-Report_Final_Links_032317.pdf.

Frisvold, D. E. (2015). Nutrition and cognitive achievement: An evaluation of the School Breakfast Program. Journal of public economics, 124, 91-104.

Glantz, F. B., Berg, R., Porcari, D., Sackoff, E., Pazer, S. (1994). School lunch eligible non-participants. Retrieved from https://naldc.nal.usda.gov/download/43357/PDF

Gleason, P.M. (1996). Student Participation in the School Nutrition Programs: An Econometric and Simulation Model. Final report submitted to the U.S. Department of Agriculture, Food and Consumer Service. Princeton, NJ: Mathematica Policy Research, Inc.

Gordon, A., Crepinsek, M. K., & Condon, E. (2007). School Nutrition Dietary Assessment Study III, Volume I: School Food Service, School Food Environment, and Meals Offered and Served (No. 2429f1634e2d4dcca2e68d5f71078b60). Mathematica Policy Research.

Gruman, D. H., Harachi, T. W., Abbott, R. D., Catalano, R. F., & Fleming, C. B. (2008). Longitudinal effects of student mobility on three dimensions of elementary school engagement. *Child development*, *79*(6), 1833-1852.

Gunderson, G. W. (2003). *The national school lunch program: Background and development*. Nova Publishers.

Ham, S., Hiemstra, S. J., & Yoon, B. J. (2002). Factors Affecting School Lunch Participation in Indiana Schools. Foodservice Research International, 13(3), 133-146.

Henry, G. T., Fortner, C. K., & Thompson, C. L. (2010). Targeted funding for educationally disadvantaged students: A regression discontinuity estimate of the impact on high school student achievement. *Educational Evaluation and Policy Analysis*, *32*(2), 183-204.

Henry, G. T. & Guthrie, J. E. (2015). Turning Around Low-Achieving Schools: A Regression Discontinuity Evaluation of North Carolina's Race to the Top-Funded School Turnaround Efforts. Conference paper. Washington, DC: Association for Education Finance and Policy 40th Annual Conference.

Henry, J. A. (2015). Analysis of the First Year Implementation of the Community Eligibility Provision of the Healthy Hunger-Free Kids Act in Kansas. University of Kansas.

Hinrichs, P. (2010). The effects of the National School Lunch Program on education and health. Journal of Policy Analysis and Management, 29(3), 479-505.

Hong, I. (2015). The Impact of the Healthy, Hunger-Free Kids Act of 2010 on National School Meal Participation Rates.

Hoyland, A., Dye, L., & Lawton, C. L. (2009). A systematic review of the effect of breakfast on the cognitive performance of children and adolescents. Nutrition research reviews, 22(02), 220-243.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, *79*(3), 933-959.

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, *142*(2), 615-635.

Kleinman, R. E., Hall, S., Green, H., Korzec-Ramirez, D., Patton, K., Pagano, M. E., & Murphy, J. M. (2002). Diet, breakfast, and academic performance in children. Annals of Nutrition and Metabolism, 46(Suppl. 1), 24-30.

Lambert, L. G., Raidl, M., Carr, D. H., Safaii, S., & Tidwell, D. K. (2007). School nutrition directors' and teachers' perceptions of the advantages, disadvantages, and barriers to participation in the school breakfast program. *J Child Nutr Manag*, *31*.

Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, *48*(2), 281-355.

Lee, V. E., & Burkam, D. T. (2003). Dropping out of high school: The role of school organization and structure. *American Educational Research Journal*, *40*(2), 353-393.

Lent, M., & Emerson, B. (2007). Preliminary findings from the 2006-2007 universal free breakfast initiative in Milwaukee Public Schools. Retrieved June, 14, 2007.

Leos-Urbel, J., Schwartz, A. E., Weinstein, M., & Corcoran, S. (2013). Not just for poor kids: The impact of universal free school breakfast on meal participation and student outcomes. Economics of education review, 36, 88-107.

Logan, C. W., Connor, P., Harvill, E. L., Harkness, J., Nisar, H., Checkoway, A., Peck, L. R., Shivji, A., Bein, E., Levin, M., Enver, A. (2014). Community Eligibility Provision Evaluation. Retrieved from https://www.fns.usda.gov/community-eligibility-provision-evaluation

Marples, C. A., & Spillman, D. M. (1995). Factors affecting students' participation in the Cincinnati public schools lunch program. Adolescence, 30(119), 745.

Maurer, K. M. (1984). The National Evaluation of School Nutrition Programs: factors affecting student participation. The American Journal of Clinical Nutrition, 40(2), 425-447.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, *142*(2), 698-714.

McDonnell, E., Probart, C., Weirich, J. E., Hartman, T., & Birkenshaw, P. (2004). School breakfast programs: perceptions and barriers. *Journal of Child Nutrition & Management*, *2*.

Mehana, M., & Reynolds, A. J. (2004). School mobility and achievement: A meta-analysis. *Children and Youth Services Review*, *26*(1), 93-119.

Meyers, A. F., Sampson, A. E., Weitzman, M., Rogers, B. L., & Kayne, H. (1989). School breakfast program and school performance. American Journal of Diseases of Children, 143(10), 1234-9.

Mirtcheva DM, Powell LM. Participation in the National School Lunch Program: Importance of school-level and neighborhood contextual factors. J Sch Health. 2009; 79: 485-494.

Moore, S. M. (2016). Principals' and School Food Service Workers' Perceptions of the Implementation of the Community Eligibility Provision.

Murphy, J. M., Pagano, M. E., Nachmani, J., Sperling, P., Kane, S., & Kleinman, R. E. (1998). The relationship of school breakfast to psychosocial and academic functioning: cross-sectional and longitudinal observations in an inner-city school sample. Archives of Pediatrics & Adolescent Medicine, 152(9), 899-907.

Neuberger, Z. (2016). House Bill Restricting Free School Meals Option Could Increase Food Insecurity in High-Poverty Neighborhoods. Retrieved from http://www.cbpp.org/research/food-assistance/house-bill-restricting-free-school-meals-option-could-increase-food

Obradović, J., Long, J. D., Cutuli, J. J., Chan, C. K., Hinz, E., Heistad, D., & Masten, A. S. (2009). Academic achievement of homeless and highly mobile children in an urban school district: Longitudinal evidence on risk, growth, and resilience. *Development and psychopathology*, *21*(2), 493-518.

Orfield, G., & Lee, C. (2005). Why segregation matters: Poverty and educational inequality.

Papay, J. P., Willett, J. B., & Murnane, R. J. (2011). Extending the regression-discontinuity approach to multiple assignment variables. *Journal of Econometrics*, *161*(2), 203-207.

PerryUndem Research (2013). California Students' Voices on School Meal Time: Results from a Survey of California Middle and High School Students.

Pollitt, E., & Mathews, R. (1998). Breakfast and cognition: an integrative summary. The American journal of clinical nutrition, 67(4), 804S-813S.

Poppendieck, J. (2010). Free for all: Fixing school food in America (Vol. 28). Univ of California Press.

Porter, K. E., Reardon, S. F., Unlu, F., Bloom, H. S., & Cimpian, J. R. (2017). Estimating causal effects of education interventions using a two-rating regression discontinuity design: Lessons from a simulation study and an application. *Journal of Research on Educational Effectiveness*, *10*(1), 138-167.

Radzikowski, J., & Gale, S. (1984). National Evaluation of School Nutrition Programs: Conclusions. American journal of clinical nutrition.

Rampersaud, G. C., Pereira, M. A., Girard, B. L., Adams, J., & Metzl, J. D. (2005). Breakfast habits, nutritional status, body weight, and academic performance in children and adolescents. Journal of the American Dietetic Association, 105(5), 743-760.

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. *Whither opportunity*, 91-116.

Reardon, S. F. (2013). The widening income achievement gap. *Educational Leadership*, *70*(8), 10-16.

Reardon, S. F., & Robinson, J. P. (2012). Regression discontinuity designs with multiple rating-score variables. *Journal of Research on Educational Effectiveness*, *5*(1), 83-104.

Reichardt, C. S., & Henry, G. T. (2012). Regression-discontinuity designs.

Richter, L. M., Rose, C., & Griesel, R. D. (1997). Cognitive and behavioural effects of a school breakfast. *South African Medical Journal*, *87*(1).

Rothstein, R. (2004). *Class and schools*. Teachers College, Columbia University.

Rumberger, R. W. (2002). Student Mobility and Academic Achievement. ERIC Digest.

Rumberger, R. W., & Larson, K. A. (1998). Student mobility and the increased risk of high school dropout. *American journal of Education*, *107*(1), 1-35.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far. *Handbook of the Economics of Education*, *3*(3), 249-277.

Sanchez, A., Hoover, L. C., Sanchez, N. F., & Miller, J. L. (1999). Measurement and evaluation of school lunch time elements in elementary, junior high, and high school levels. Journal of child nutrition & management.

Schanzenbach, D. W., & Zaki, M. (2014). Expanding the school breakfast program: Impacts on children's consumption, nutrition and health (No. w20308). National Bureau of Economic Research.

Schochet, P., Cook, T., Deke, J., Imbens, G., Lockwood, J. R., Porter, J., & Smith, J. (2010). Standards for Regression Discontinuity Designs. *What Works Clearinghouse*.

Segel, B., Hewins, J., Sanderson, M., Nchako, C., Neuberger, Z., Cai, L., Maurice, A. (2016). Community Eligibility Adoption Rises for the 2015-2016 School Year, Increasing Access to School Meals. Retrieved from http://www.cbpp.org/research/food-assistance/community-eligibility-adoption-rises-for-the-2015-2016-school-year

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Wadsworth Cengage learning.

Somers, M. A., Zhu, P., Jacob, R., & Bloom, H. (2013). The Validity and Precision of the Comparative Interrupted Time Series Design and the Difference-in-Difference Design in Educational Evaluation. *MDRC*.

Staiger, D. O., & Stock, J. H. (1994). Instrumental variables regression with weak instruments.

St. Clair, T., Hallberg, K., & Cook, T. D. (2016). The validity and precision of the comparative interrupted time-series design: three within-study comparisons. *Journal of Educational and Behavioral Statistics*, *41*(3), 269-299.

Taras, H. (2005). Nutrition and student performance at school. Journal of school health, 75(6), 199-213.

Tucker, C. J., Marx, J., & Long, L. (1998). " Moving on": Residential mobility and children's school lives. *Sociology of Education*, 111-129.

United State Department of Agriculture [USDA]. (2015). The Community Eligibility Provision (CEP): What Does It Mean For Your School or Local Educational Agency? Retrieved from https://www.fns.usda.gov/sites/default/files/cn/CEPfactsheet.pdf

van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression discontinuity approach. International Economic Review, 43(4), 1249–1287.

Wahlstrom, K. L., & Begalle, M. S. (1999). More Than Test Scores: Results of the Universal School Breakfast Pilot in Minnesota. Topics in clinical nutrition, 15(1), 17-29.

Wesnes, K. A., Pincock, C., Richardson, D., Helm, G., & Hails, S. (2003). Breakfast reduces declines in attention and memory over the morning in schoolchildren. Appetite, 41(3), 329-331.

Westervelt, E. (2013). These Days, School Lunch Hours Are More Like 15 Minutes. *Nashville Public Radio*. Retrieved from http://www.npr.org/templates/transcript/transcript.php?storyId=248511038

What Works Clearinghouse (WWC). (2014). Procedures and Standards Handbook, Version 3.0. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf

Wong, V. C., Steiner, P. M., & Cook, T. D. (2013). Analyzing regression-discontinuity designs with multiple assignment variables: A comparative study of four estimation methods. *Journal of Educational and Behavioral Statistics*, *38*(2), 107-141.

Worobey, J., & Worobey, H. S. (1999). The impact of a two-year school breakfast program for preschool-aged children on their nutrient intake and pre-academic performance. Child Study Journal, 29(2), 113-113.

Zellner, A. B. (2016). High School Students' Participation and Satisfaction of the School Lunch Program. *Mount Mary University*.

Zimmer, R., Henry, G. & Kho, A. (2017) The Effects of School Turnaround in Tennessee's Achievement School District and Innovation Zones. *Educational Evaluation and Policy Analysis*, 0162373717705729.

Zimmerman, D. J. (2003). Peer effects in academic outcomes: Evidence from a natural experiment. *The Review of Economics and statistics*, *85*(1), 9-23.

CHAPTER 2

THE SCHOOLS TURNAROUND LEAVES BEHIND: SYSTEMIC EFFECTS OF RECRUITING HIGH-PERFORMING TEACHERS FOR SCHOOL TURNAROUND

Introduction

In the early 21st century, No Child Left Behind (NCLB) shone a spotlight on the nation's lowest-performing schools, pressuring many schools, districts, and states to implement school and district turnaround reforms aimed at improving the performance of these schools. Through School Improvement Grants and Race to the Top funding, the federal government provided states, districts, and schools with funds to facilitate these turnaround initiatives, which has led to reforms ranging from intrusive state takeover to more collaborative partnerships between local education agencies and individual schools. A large number of these turnaround models have relied on a key element in their theory of change: recruiting and hiring of high-performing teachers (USDOE, 2009; 2010). For example, many states and districts have attracted high-performing teachers to difficult-to-teach environments by offering recruitment, retention and performance incentives. While many low-performing schools have benefitted from the presence and work of these teachers (Zimmer, Henry & Kho, 2017), little attention has been given to an unintended consequence of these turnaround efforts – the effects on schools those high-performing teachers left, henceforth, referred to as "sending schools".

In this analysis, we utilize a statewide, student-level, longitudinal dataset to study the unintended consequences of teacher recruitment into the Memphis Innovation Zone (iZone). The Memphis iZone has been regarded as one of the most successful turnaround initiatives aimed at raising student achievement (Gonzales, 2016; Kebede, 2016; Tillery, 2017; Zimmer, Henry & Kho, 2017), and a prominent strategy for iZone schools has been to hire highly effective teachers ("iZone," 2017). Zimmer, Henry & Kho (2017; Henry et al., 2017) show iZone schools have successfully recruited effective teachers; however, many of those teachers came from nearby districts and even schools within the same district. In this analysis, we ask: To what extent has Memphis iZone schools' practice of recruiting high quality teachers affected the achievement of students in the sending schools? Using value-added measures of student performance in a series of fixed effect models that follow recent research on the effects of teacher turnover (Ronfeldt, Loeb & Wyckoff, 2013; Henry & Redding, 2017), we examine the changes in student test score gains in sending schools after the teachers left for an iZone school. The results of this analysis have direct implications for the short-run unintended consequences of incentivizing teacher transfers and better understanding the supply of effective teachers within an urban school district. To be even more explicit, if one of the consequences of successfully recruiting highly effective teachers into low-performing school is reducing student achievement gains in the schools from which they were recruited, the supply of effective teachers may have a ceiling effect and the teacher labor market, at least in the short run, may function as a zero-sum game.

In the next section, we draw upon the teacher incentive and teacher turnover literature to inform the discussion on systemic effects of recruiting teachers for school turnaround programs. We then discuss the turnaround initiative implemented in Tennessee and the hiring and recruitment practices of the Memphis iZone, which leads to the research questions for this study. Next is a description of the data and the methods used, followed by the results of our analysis. We conclude with a discussion of our findings and suggest future research.

Literature Review

The literature on competition in schools typically highlights school choice markets in which students have the flexibility to choose the schools they attend. Schools must compete with one another to attract and recruit students in order to remain in operation. In doing so, school choice advocates hope that the competition for students would motivate all schools to improve their performance (Hoxby, 2000; Sass, 2006; Zimmer and Buddin, 2009). There is currently, however, a void in the literature of understanding how schools compete for one of the key ingredients in raising student achievement – teachers. Extant literature is clear that teacher quality matters. Students taught by more effective teachers, as determined by various value-added measures, have higher test score gains, more positive non-cognitive outcomes (such as school attendance and behavior), and better long-term outcomes, including being less likely to have teenage pregnancies, more likely to attend college, and earn higher salaries (Sanders, Wright & Horn, 1997; Rockoff, 2004; Rivkin, Hanushek & Kain, 2005; Aaronson, Barrow & Sander, 2007; Koedel & Betts, 2007; Hanushek, 2011; Jackson, 2012; Chetty, Friedman & Rockoff, 2014; Jackson, Rockoff & Staiger, 2014). Attracting, recruiting, and retaining high quality teachers, therefore, is indispensable for schools to produce high levels of educational outcomes. However, in contrast to school competition for students which only occurs in markets with options for school choice, competition for teachers occurs in all school markets since teachers have the option to apply to and work in any school, provided they meet the necessary qualifications. This is particularly true for high quality teachers who are, therefore, more attractive in the teacher labor market.

Unfortunately, research shows that highly effective teachers are less likely to work in schools with primarily underserved minorities and lower-performing students. Steele and colleagues (2015) find that in one southern, large, urban school district, students in schools in the highest quartile of minority enrollments have teachers with value-added estimates that are 0.11 standard deviations lower than those in schools in the lowest minority quartile. This pattern extends to students in high-poverty elementary and middle schools throughout the nation (Glazerman & Max, 2011; Sass et al., 2012; Isenberg et al., 2013). Examining several measures of teacher quality, a study in Washington state by Goldhaber, Lavery & Theobald (2015) also found inequitable distributions – students in schools with high percentages of free and reduced-price lunch status, underrepresented minorities, and/or low prior academic performance had teachers with lower value-added scores, years of experience, and licensure exam scores. Together, these studies show that, left to their own devices, more highly effective teachers will naturally sort themselves away from schools with students who need them the most.

A growing body of research provides credible evidence that financial incentives for effective teachers to work in high poverty, high minority, and low performing schools does increase the effectiveness of teachers in those schools but findings about the retention of those teachers are mixed. Two recent studies (Steele, Murnane & Willett, 2010; Cowan & Goldhaber, 2015) showed that bonuses can attract teachers whose characteristics are associated with more effective teachers into lower performing schools, but the bonuses did not affect retention of those teachers in the more challenging schools. Another study of an incentive program that offered bonuses to attract effective teachers into low-performing schools by Glazerman and colleagues (2013) showed positive effects on attracting those teachers, their retention in the schools and student achievement. Two additional studies have shown that retention bonuses for effective teachers in low performing schools have had positive effects on teacher retention (Clotfelter et al. 2008; Springer, Swain & Rodriguez, 2016). These prior studies have identified a plausible causal impact of teacher recruitment incentives into low-performing and low-income schools on teacher recruitment, teacher retention, and student achievement in the recruiting school, but there have been no studies, to our knowledge, that address any systemic effects that may result from this recruitment. This study seeks to fill that void.

Redding and Henry (2017) highlight three mechanisms through which teacher turnover can impact student achievement. When teachers leave a school, staff instability severs working relationships between those teachers and administrators and other teachers, and connections built with students are lost. A second mechanism, classroom disruptions, occurs when teachers leave through mid-year transitions and students have to acclimate to a different teacher and a new set of routines. Lastly, teacher turnover changes the composition of teachers in a school. If a more effective teacher is replaced by a less effective teacher, student academic performance will suffer.

Previous research provides evidence that teacher turnover negatively impacts student achievement. Two studies (Ronfeldt, Loeb & Wyckoff, 2013; Hanushek, Rivkin & Schiman, 2016) find that students in grades in which teachers turned over prior to the school year starting have reduced test score gains from 4 to 11 percent of a standard deviation on average. Henry & Redding (2017) find less substantial and inconsistent negative impacts when turnover during the prior summer and within the school year are combined. However, when disaggregating this by within-year and end-of-year turnover, the test score gains of students of a teacher who departs during the school year are reduced by 5 to 12 percent of a standard deviation but approach zero and even become positive in some cases of end-of-year moves. Lastly, prior literature also suggests that teachers who leave tend to be less effective than those who stay (Hanushek, Kain & Rivkin, 2004; Goldhaber, Gross & Player, 2007; Boyd et al., 2005; 2008a; 2008b; 2011; Henry, Bastian & Fortner 2011; Henry, Fortner & Bastian, 2012; Hanushek, Rivkin & Schiman, 2016).

In this study, we investigate the impact of highly effective teachers leaving schools as a result of being recruited to teach in a low-performing school in a potentially zero-sum teacher labor market. To the extent that losing a highly effective teacher is more likely to influence each of the three mechanisms discussed previously than losing a less effective teacher, we would expect the impact estimates of teacher turnover in prior studies to underestimate the impact of when highly effective teachers leave a school. Because highly effective teachers are often tapped to serve in leadership roles, these teachers may have more institutional knowledge and may more often facilitate collaborative efforts. To the extent that this is true, the loss of a highly effective teacher would be more detrimental than the loss of a less effective teacher. Prior literature also shows that highly effective teachers are more likely to have a set of routines and procedures for their classrooms (Emmer & Stough, 2001; Oliver & Reschly, 2007). Therefore, losing a highly effective teacher can also create greater classroom disruption. Lastly, differences in quality of replacement and replaced teacher is most clearly affected by the difference in leaving teacher quality as this difference would be greater if the replaced teacher was more effective. In summary, drawing these teachers from other schools may yield positive effects at the schools they are recruited to, but may also lead to negative effects from turnover at the schools these teachers leave.

Lastly, it is important to investigate the characteristics of the school from which teachers are drawn. Previous analyses of heterogeneous effects by school characteristics have found that teacher turnover is more harmful for lower-achieving and highly-economically disadvantaged schools (Ronfeldt, Loeb & Wyckoff, 2013; Hanushek, Rivkin & Schiman, 2016). Therefore, if any of these highly effective teachers are being drawn from other low-performing schools in which these teachers may serve as mentors or hold leadership positions, the unintended consequences of the teacher recruitment strategies could be even more harmful to the students in the schools those teachers left.

The Tennessee Context

In 2011, Tennessee, like many states, applied for and was awarded a waiver from the seemingly unattainable NCLB goal of having 100% of students proficient in reading and math by 2014. As a part of that waiver, the state agreed to identify its lowest-performing 5% of schools, labeling them Priority schools. These schools resided primarily in the largest cities of the state – 69 in Memphis, six in Nashville, six in Chattanooga, and two in two smaller school districts. In addition to labeling, the state decided that each of the Priority schools would be subject to one of three interventions to improve their status. Through aggressive interventions, such as district-within-a-district Innovation Zones, Tennessee signaled that it was prepared to take on bold initiatives to turnaround these schools.

Under the Race to the Top grant and previous School Improvement Grant guidelines, the federal government required states to choose one of four reform models to turnaround schools – transformation, turnaround, restart, and closure. The transformation model required replacement of the principal, increased learning time, more rigorous teacher evaluation systems, and additional autonomy for schools, freeing them of district bureaucracy. The turnaround model took reform to the next level, requiring all of the changes of the transformation model but also replacement of at least 50% of the school staff. The restart model required the transfer of school management to a separate entity such as a CMO. Under this model, the majority, if not all, of the school staff would be replaced under a new manager and management system. The last model, closure, closed the low-performing school. Among previous School Improvement Grants, most schools chose the transformation model – the least intrusive of them all (Dragoset et al., 2017). However, Tennessee's First to the Top legislation in 2010 and Race to the Top application in 2011 outlining its turnaround initiatives highlighted turnaround and restart reform models, proving that the state was prepared to confront the status quo in which these schools had been allowed to languish in the lowest rungs of performance and engage in major reforms, and earning them one of only two Race to the Top grants awarded in phase one of the federal program.

The intervention encompassing the largest number of Priority schools followed the federal turnaround model. Districts with large numbers of Priority schools could open a district-within-a-district that would focus on school turnaround, called an Innovation Zone (iZone). The schools in these iZones would remain under the operation of the larger district but would receive new leadership and increased autonomy from the larger district. The three large cities housing the majority of Priority schools – Memphis, Nashville, and Chattanooga – all opened iZones. In the first year of Priority status, 2012-13, Nashville opened its iZone with four schools and Memphis opened its iZone with seven schools. The following year, Memphis added six more schools to their iZone, and Chattanooga began their iZone with five Priority schools. In the 2014-15 and 2015-16 school years, Memphis added four and one more school, respectively.[7]

One of the key strategies for turning around these low-performing schools was recruiting and retaining highly effective teachers (USDOE 2009; 2010). To assess teacher effectiveness, schools utilized the Tennessee Educator Assessment Model (TEAM), the state's teacher evaluation program. Tennessee's teachers are rated each year through both qualitative and quantitative measures, including classroom observations, individual conferences, student growth

[7] The Tennessee Department of Education released a new list of Priority schools in 2014, which allowed Knoxville to open an iZone as well. All but one of their Priority schools were not on the original Priority list.

on the state standardized assessments, and other school-based student achievement measures. Together, scores from each of these components form an overall level of effectiveness (LOE) score for each teacher, ranging from 1 to 5. Table 1 shows the distribution of statewide LOE scores for the 2014-15 school year.

Table 1. Distribution of Teacher Level of Effectiveness Scores in 2014-15

Level of Overall Effectiveness	Description	Percent Receiving Score
1	Significantly Below Expectations	4%
2	Below Expectations	12%
3	At Expectations	26%
4	Above Expectations	32%
5	Significantly Above Expectations	26%

To attract high quality teachers into the state's lowest-performing schools, the Tennessee Department of Education offered signing and retention bonuses to teachers with proven effectiveness. Teachers who were rated level 5 in the previous year, the highest possible rating on the Tennessee teacher evaluation system, were eligible for a $7,000 signing bonus if they committed to working at a Priority school for at least two years (TDOE, 2013). Teachers who were rated level 5 in the previous year and were already teaching at a Priority school were offered $5,000 to continue working at a Priority school (TDOE, 2013; Springer et al., 2016). The Memphis iZone also offered $1,000 to $1,500 bonuses to Level 4 and 5 teachers who agreed to teach for three years with an additional $1,000 annual award if teachers met district benchmarks (Sullivan, 2013; Burnette, 2017). In addition, as of 2015-16, all Memphis teachers who scored a level 5, 4, or 3 rating received a $1,200, $1,000, or $800 increase in salary as well (USDOE, 2016).

Springer, Swain & Rodriguez (2016) evaluated the state's signing bonus and found that teachers in tested subjects and grades were 20% more likely to stay as a result of the bonus. While the effects of the other incentives have yet to be evaluated, descriptive analyses of the distribution of teacher quality and teacher mobility in the Achievement School District (ASD), the state-run school district with authority to takeover and restart Priority schools, and iZones have shown an increase in the hiring and retention of highly effective teachers, as determined by value-added measures. Zimmer, Henry & Kho (2017) found that in the first three years of implementation, both ASD and iZone schools did a better job of hiring more highly effective teachers than other Priority schools and other Tennessee schools throughout the state. Using Tennessee's Teacher Value-Added Assessment System (TVAAS) score, an annual rating between 1 (least effective) and 5 (most effective) of teacher performance based on student-level value-added growth scores that serves as one component of the LOE score described previously, teachers hired into the ASD scored on average 3.34 and those hired into the iZone scored on average 3.37 over their first three years of implementation. This is in comparison to other Priority schools, 2.80, and other non-Priority Tennessee schools, 3.18, in those same years. The iZone schools also did a better job at retaining more highly effective teachers with an average TVAAS score of 3.43 than both the ASD, 2.97, and other Priority schools, 2.95. (The iZone performed comparably with other non-Priority Tennessee schools.) In a subsequent study, Henry and colleagues (2017) confirm these teacher effectiveness patterns but also found that iZones not only effectively hired high quality teachers, but also were able to develop other retained teachers into the highest-performing category during their tenure at iZone schools.

The authors note one caveat to these teacher performance ratings – TVAAS scores are only available for teachers who taught in a tested grade or subject in a Tennessee public school in the year prior. iZone schools replaced a much smaller portion of their staff with novice teachers or teachers new to the Tennessee public school system than the ASD. In 2012-13, the first year of implementation, 19% of the staff in iZones fell in this category, compared to 31% in ASD schools run directly by the state and 66% in ASD schools contracted out to CMOs (Henry et al., 2015). Therefore, the TVAAS scores above represent a larger proportion of the teachers in iZones than in the ASD.

As Zimmer, Henry & Kho (2017) state, this ability to effectively recruit and retain high quality teachers may have been an important reason iZone schools were so effective at raising student achievement. In their evaluation of the ASD and iZones, Zimmer, Henry & Kho found that after three years of implementation, schools in iZones yielded, on average, student test score gains of 0.10 to 0.20 standard deviations larger than other Priority schools, depending on subject. Though positive effects were found in all three iZones, the strongest and most consistent effects were in Memphis.

The question we investigate in this paper is whether these positive results for Memphis iZone schools came at the expense of other schools. In a zero-sum teacher labor market, the number of high quality teachers is fixed. Therefore, if the positive results for Memphis iZone schools are driven by their recruitment of highly effective teachers, the limited number of available highly effective teachers in the pool remaining necessarily means a loss in the quality of teachers at the schools they leave. This consequence of teacher turnover, both in general and particularly of higher quality teachers in the case of the sending schools, can be expected based on prior research to translate into poorer student achievement at sending schools. However, it may be possible that the schools who lose effective teachers have natural advantages in subsequently recruiting effective teachers and the quality of teachers remains high. This would not be expected for the highly segregated, high poverty Priority Schools that lose effective teachers to an iZone.

Across all three cohorts of Memphis iZone schools, about 505 teachers transferred into a Memphis iZone school within the first three years of iZone status, 181 of whom taught a tested subject or grade. Of these transferring teachers, 93% moved from within the school district; 5% of these moved from a bordering or nearby school district; and 2% moved from other schools throughout the state.

While these high-quality teachers were encouraging additions for the Memphis iZone schools, the other side of the potential zero-sum story remains. Many other schools had to lose their highly effective teachers in order for the Memphis iZone schools to gain them. The 181 transferring teachers came from 100 different schools, averaging a loss of nearly two teachers per school, though several schools lost as many as six teachers in one year and one up to 14 teachers over the three-year period. Thirty-three of the schools that lost teachers were themselves Priority schools, meaning many schools that were already struggling to improve lost some of their best talent. Again, assuming a fixed supply of highly effective teachers that are also willing to work in a low-performing school in Memphis, these leaving teachers are usually replaced with less effective teachers (Darling-Hammond & Sykes, 2003; Ronfeldt, Loeb & Wyckoff, 2013), lowering the overall effectiveness of teachers remaining at these schools. Even the best-case scenario in which a leaving teacher is replaced with a highly effective teacher who has previously worked in the school but in another grade can have negative consequences. Atteberry and colleagues (2017) investigate the impact of different types of teacher churn on student achievement and find that effects of teacher churn are negative, regardless of teacher quality and from where the replacing teacher came.

Anecdotal evidence suggests that schools that lost teachers who were recruited to the Memphis iZone have suffered losses in state assessments scores (Williams, 2016). However, the effects of this teacher loss have not yet been estimated. This study seeks to investigate the zero-sum teacher labor market assumption and provide a causal impact of the teacher turnover created by recruitment of highly effective teachers into the Memphis iZone on the students in the sending schools. The data we use allows us to identify the effect with grade-level granularity and to examine differential effects based on school-level characteristics. Specifically, we ask, to what extent has Memphis iZone schools' practice of recruiting high quality teachers affected the achievement level of schools and students in the schools from which they came? Further, does this effect vary by the proportion of economically disadvantaged students in the school or the school's prior performance status? To our knowledge, this will be the first study to examine the unintended consequences of a program that recruits highly effective teachers to turnaround their lowest-performing schools. As states and districts continue to address these schools, it is important to understand the implications of one of the most commonly used turnaround practices – recruiting highly effective teachers – on the schools most potentially negatively impacted by this practice.

Methods

Data & Measures

This study utilizes 2011-12 to 2014-15 statewide administrative datasets from the Tennessee Department of Education (TDOE) and managed by the Tennessee Education Research Alliance (TERA) at Vanderbilt University. The first dataset is student-level and includes demographic data, standardized test scores, and school and grade assignment variables for all students in the Tennessee public school system for each school year. A second dataset is teacher-level and includes school assignments and grade(s) and course(s) taught.

The key dependent variables in this analysis are student test scores. In Tennessee, all students in grades three to eight are tested on an annual basis in reading, math, and science using the Tennessee Comprehensive Assessment Program (TCAP). Students in high school are tested at the end of a select group of courses (English I, English II, English III, Algebra I, Geometry, Algebra II, Biology, and Chemistry). For this analysis, TCAP test scores will be standardized by year and grade; End of Course (EOC) test scores will be standardized by year and subject.

In the dataset, we can track teacher school assignments from year to year. Further, we can identify the grade(s) and subject(s) elementary and middle school teachers taught and the course(s) high school teachers taught. This allows us to create a continuous variable that identifies the proportion of teachers exiting grade g in school s in year $t-1$ to enter a Memphis iZone school. This teacher turnover proportion will serve as the main independent variable.

To account for several factors correlated with both teacher transfer and student achievement, we include a series of covariates at the student and school levels. At the student level, these covariates include gender, race, free and reduced-price meal status, special education status, English language learner status, and mobility status, a binary indicator of whether the student was new to the school in the given year. We also include student's prior year test scores for a value-added interpretation of the effect estimates. To account for school-level differences, we aggregate student-level data up to the school-level to include the percent of students that are economically disadvantaged, the percent of students of racial/ethnic minority status, and the percent of students that were mobile that year. We also control for school level (elementary, middle, or high).

In addition to teachers leaving for Memphis iZone schools, there may be other teachers that exit these schools for other reasons. To avoid misattributing the effects of other turnover to Memphis iZones, we include a teacher turnover control variable that excludes moves to iZones. By including this other teacher turnover variable, we can also assess whether the impact of losing teachers to the Memphis iZone schools, who are disproportionately highly effective teachers, is greater than typical teacher turnover.

One important limitation to note is with respect to the operationalization of the teacher turnover independent variables. Our data does not allow us the ability to distinguish between the different reasons teachers leave their schools. Therefore, we cannot identify if teachers who left for the iZone would have left regardless of the iZone opportunity. For this analysis, we assume that teachers would not have left their schools had the iZone opportunity not been available. But to the extent that this group of teachers exists and were higher-performing, our main effect estimates would be biased upward.

Empirical Framework & Samples

The ideal method for evaluating the effect of the Memphis iZone teacher recruitment is an experiment that randomly assigns teachers to transfer and not to transfer from their schools to the Memphis iZone. The difference in achievement of the treatment group, students assigned to the grades/courses from which teachers transfer, and the comparison group, students assigned to the grades/courses from which teachers do not transfer, would yield the impact of the recruitment strategy. However, this experiment is infeasible as we cannot assign teachers to work at specific schools. We, therefore, utilize what we believe to be the next best method to provide a causal impact of the Memphis iZone recruitment on the schools and students the teachers left – a series of value-added equations to estimate student achievement gains along with student-, school-, year- and grade-level fixed effects, similar to those used in two recent studies (Ronfeldt, Loeb & Wyckoff, 2013; Henry & Redding, 2017).

In particular, we are concerned with selection bias of schools that "receive treatment." In other words, the types of schools these teachers leave are likely different from other schools that did not lose teachers to the Memphis iZone. The schools could have unobserved school climates or neighborhood characteristics that affect their student achievement. To address this selection issue, our primary model utilizes a school-by-year fixed effect, which controls for time-invariant observed and unobserved school-by-year characteristics. We prefer a school-by-year fixed effect over a school fixed effect to leverage variation in turnover by grade within the same school within the same year. The preferred specification allows us to control for possible temporal shocks that affect both teacher turnover and student achievement. For instance, principal turnover at a school in one particular year may simultaneously influence both student achievement and teacher turnover, biasing the estimate of the effects of turnover on achievement. By including a school-by-year fixed effect, we can control for this principal turnover and other omitted variables specific to the school and year. We model this approach as:

$$y_{igst} = \beta_0 + \beta_1\, iZoneTchrGradeTurnover_{gst-1} + \beta_2\, OtherTchrGradeTurnover_{gst-1} + \beta_3\, y_{igst-1} + S_{igst}\, \mathrm{B}_j + \gamma_{st} + e_{igst} \quad (1)$$

where y represents the test score for student i in grade g, school s, and year t. *iZoneTchrGradeTurnover* is a continuous variable indicating the proportion of teachers having left grade g in school s in year $t - 1$, the year prior to a student entering the respective grade, to teach at a Memphis iZone school. β_1 is the key coefficient of interest and identifies the effect of 100% grade-level teacher turnover as a result of losing teachers to the Memphis iZone. $OtherTchrGradeTurnover_{gst-1}$ is a continuous variable indicating the proportion of teachers having left grade g in school s in year $t - 1$ that did not move to a Memphis iZone school. This variable ensures that the effects of grade-level turnover for reasons other than leaving to join a Memphis iZone are not erroneously attributed to the effects of leaving for Memphis iZone schools. Additionally, because the recruitment of teachers to Memphis iZone schools specifically targeted high-performing teachers and the loss of a high-performing teacher is more harmful than losing a lower-performing teacher, we expect β_1 to be negative and larger than β_2. y_{igst-1} represents the student's test score in the year prior, S_{igst} represents a vector of student characteristics (gender, race, economically disadvantaged status, special education status, English language learner status, and mobility status), γ_{st} represents the school-by-year fixed effect, and e_{igst} is an idiosyncratic error term. Standard errors are clustered at the school level.

This school-by-year fixed effect model will allow us to account for any time-varying confounding changes in schools by comparing the effects of teacher turnover that occurred in one grade due to a teacher transferring to an iZone school to other grades in the same school and year that did not lose teachers to the iZone, adjusting for any other teacher turnover at the school in the prior year. Tested students attending schools in any year following a year when the school lost a teacher to the Memphis iZone (regardless of grade) are included in this analytic sample.

This first model, however, could be biased by within-school grade-level differences. For example, a teacher may choose to transfer out of a school due to lack of effort on the part of her grade-level peers that other grades may not experience. Therefore, we estimate a second model by replacing the school-by-year fixed effect with a school-by-grade fixed effect, which allows us to control for the characteristics of her peers that remain constant and other omitted variables specific to the grade and school and exploit the variance in turnover within the same grade and school but over time:

$$y_{igst} = \beta_0 + \beta_1\, iZoneTchrGradeTurnover_{gst-1} + \beta_2\, OtherTchrGradeTurnover_{gst-1} + \beta_3\, y_{igst-1} + S_{igst}\, \mathrm{B_j} + X_{st}\, \mathrm{B_k} + \delta_{gs} + \theta_t + e_{igst} \qquad (2)$$

In this specification, the within-school differences in student achievement gains before and after teachers transferred to the iZone are used to estimate the effects from losing a teacher to the iZone, which restricts the analytical sample to tested students enrolled in any school-grade combination in which a teacher, in any year, left for an iZone school. In addition to student characteristics, we control for school-level characteristics (percent economically disadvantaged, percent minority, and percent mobile) for school s at time t, which is represented by X_{st}, and employ a year fixed effect θ_t to adjust for overall yearly differences. δ_{gs} represents the school-by-grade fixed effect.

Both the school-by-grade and school-by-year fixed effects models fail to account for the bias due to the possibility of nonrandom student sorting into classes with higher-performing teachers who were recruited to the iZone schools. To account for this and other unobserved student characteristics, we estimate a third model using student fixed effects:

$$y_{igst} = \beta_0 + \beta_1\, iZoneTchrGradeTurnover_{gst-1} + \beta_2\, OtherTchrGradeTurnover_{gst-1} + X_{st}\, \mathrm{B_k} + \phi_i + \theta_t + e_{igst} \qquad (3)$$

We again include school-level control variables and a year fixed effect. ϕ_i represents the student fixed effect. This specification compares students' performance in years in which they were in a grade which experienced teacher turnover to a Memphis iZone school the previous year to their own performance in years in which they did not experience this turnover. The effect of losing a teacher to an iZone school is identified for the sample that includes any tested students who have ever entered a grade in which the teacher exited to teach at a Memphis iZone school and has test scores before and after the teacher's move.

By including all three approaches, we can examine whether our results are robust to the assumptions of the various models. However, we identify the school-by-year fixed effect approach (equation 1) as our preferred model as we are most concerned with selection bias of schools (rather than students), and we believe the most influential factors that simultaneously affect both student achievement and teacher turnover are likely to occur at the school level in a particular year than consistently each year and differentially across grades.

Finally, we extend these analyses by investigating whether the effects of teacher turnover to Memphis iZone schools has heterogeneous effects based on prior school-level characteristics – percent economically disadvantaged and prior performance level. Previous literature has found that teacher turnover is more harmful for students in more economically disadvantaged schools and lower-performing schools (Ronfeldt, Loeb & Wyckoff, 2013; Hanushek, Rivkin & Schiman, 2016). We test whether highly effective teacher turnover is more harmful for students in schools with these characteristics by interacting these moderating variables with $iZoneTchrGradeTurnover_{gst-1}$. For the school percentage of economically disadvantage, we compare the upper and lower quartiles of schools to the middle half. For school performance level, we use an indicator identifying sending schools that are Priority schools (the state's lowest performing 5% of schools).

Results

Overall Results

In Table 2, we first compare the baseline school-level characteristics of sending schools and iZone schools. We also include the characteristics of all schools in Memphis as a reference[8].

[8] We exclude alternative schools.

iZone schools were primarily elementary and middle schools. Therefore, most sending schools were also elementary and middle schools. Sending schools had smaller percentages of minority and economically disadvantaged students than iZone schools – 91% minority and 83% economically disadvantaged in sending schools compared to 99% minority and 91% economically disadvantaged in iZone schools – but greater than the average Memphis school, which was 86% minority and 74% economically disadvantaged. Similarly, the sending schools, which scored 0.50 to 0.75 standard deviations below the state average depending on subject, were higher-performing on the state's standardized assessments than iZone schools, which scored 0.92 to 1.20 standard deviations below average, but worse than the average Memphis school, which scored 0.26 to 0.34 standard deviations below average. In addition, one-third of sending schools were also Priority schools (the state's lowest-performing 5% of schools). Overall, the statistics in table 2 show that the schools teachers left were not, on average, low-minority, high-performing schools. Teachers that transferred to the Memphis iZone left only slightly more disadvantaged and lower-performing teaching environments than the iZone schools they entered.

Table 2. Baseline School-Level Characteristics of Sending Schools, Memphis iZone Schools, and Memphis Schools

Characteristics	Sending Schools	iZone Schools	Memphis
Total Schools	100	17	1,614
Elementary	40%	47%	59%
Middle	39%	35%	24%
High	21%	18%	17%
Percent Minority	91%	99%	86%
Percent Economically Disadvantaged	83%	91%	74%
Priority Schools (Lowest-Performing)	33%	100%	28%
Reading Score*	-0.65	-1.10	-0.27
Math Score*	-0.50	-0.92	-0.26
Science Score*	-0.75	-1.20	-0.34

Alternative schools are excluded.
*Test scores represent average standardized test scores in years prior to teacher recruitment/loss and are standardized at the state level.

In Table 3, we display the estimated effects of teachers leaving for the Memphis iZone on student test scores of the grades and subjects in the sending schools in the year after teachers leave. Columns 1, 4, and 7 provide the results of our preferred model – the school-by-year fixed effect model; columns 2, 5, and 8 provide the results of the school-by-grade fixed effect model; and columns 3, 6, and 9 provide the results of the student fixed effect model for reading, math, and science, respectively. For each model, we also display the coefficient for other teacher turnover as a comparison and indicate in bold statistically different coefficient estimates relative to the coefficient estimates for teacher turnover to the Memphis iZone[9]. Each of the coefficient estimates should be interpreted as the change in test score gains for students entering a grade in which all teachers left the previous year. On average, grades that lost reading, math, and science teachers to the Memphis iZone lost 57%, 66%, and 67% of their grade-level teachers, respectively.

[9] We test this at the 95% confidence level.

Our preferred model shows that students in grades that lost 100% of their reading teachers to the Memphis iZone scored 0.10 standard deviations lower on their reading assessment than students in the same school in the same year that did not lose any teachers to the Memphis iZone. We did not find any positive or negative effects of losing reading teachers to other reasons in this model. The student fixed effect model in reading also yields a comparable negative effect of teacher turnover to the Memphis iZone – students in grades that lost 100% of their reading teachers to the Memphis iZone scored 0.081 standard deviations lower on their reading assessments. The effect of losing teachers to reasons other than the Memphis iZone is -0.057 standard deviations. While greater in magnitude in both cases, the effect estimate for teacher turnover to iZone is not statistically different from the effect estimate for other teacher turnover in either case.

In math, we do not find effects in our preferred model. The standard errors in both the school-by-year and school-by-grade fixed effect models are substantially larger (than reading) which renders these effects as statistically indistinguishable from zero. In the student fixed effect model, however, we find that students in grades that lost 100% of their math teachers to the Memphis iZone scored 0.197 standard deviations lower than they did in years in which none of their grade-level math teachers left for the iZone, which is approximately 0.08 standard deviations less than the effect of losing teachers to other reasons. This difference is statistically significant.

In science, we again do not find effects in our preferred model. However, the school-by-year and student fixed effect models both yield negative effects. Students entering grades that lost 100% of their science teachers to the Memphis iZone scored 0.17 to 0.20 standard deviations lower than the respective comparison groups. This is statistically different from the effect of other teacher turnover only in the student fixed effect model.

Table 3. Estimates of the Effects of Teacher Turnover to Memphis iZone on Student Achievement

Variable	Reading			Math			Science		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Teacher Turnover to iZone	-0.100*	-0.062	-0.081***	-0.121	-0.148	-0.197***	-0.084	-0.171*	-0.199***
	(0.041)	(0.045)	(0.017)	(0.116)	(0.076)	(0.018)	(0.053)	(0.078)	(0.020)
Other Teacher Turnover	-0.009	0.009	-0.057**	-0.156	**0.102**	**-0.279*****	**0.131**	-0.051	**0.047**
	(0.059)	(0.059)	(0.019)	(0.134)	**(0.097)**	**(0.030)**	**(0.092)**	(0.099)	**(0.035)**
School x Year FE	X			X			X		
School x Grade FE		X			X			X	
Student FE			X			X			X
R squared	0.630	0.617	0.052	0.474	0.477	0.068	0.540	0.515	0.030
N	13,043	16,087	18,753	9,839	13,202	15,081	10,260	14,192	13,620

* p<0.05, ** p<0.01, *** p<0.001
Student Controls: Gender, Race, FRPL status, Special Education status, ELL status, Mobility Status, Prior Reading Test Score, Prior Math Test Score, Prior Science Test Score
School Controls: Percent Minority, Percent FRPL, Percent Student Mobility, School Level

Heterogeneous Effects

Previous literature (Ronfeldt, Loeb & Wyckoff, 2013; Hanushek, Rivkin & Schiman, 2016) has found that teacher turnover is more harmful for students in more economically disadvantaged schools and lower-performing schools. We test these hypotheses using our preferred model (school-by-year fixed effect) in table 4. The first three columns display the effect estimates of losing teachers to the Memphis iZone based on the percentage of students that are economically disadvantaged in the school. We compare schools in the top quartile (most economically disadvantaged) and bottom quartile (least economically disadvantaged) to the middle half of economically disadvantaged schools (our omitted group). In reading and math, our estimates are statistically indistinguishable from zero. (No math teachers exited the lowest economically disadvantaged quartile of schools.) In science, students attending schools in the middle half of socioeconomic status scored 0.19 standard deviations greater as a result of losing 100% of grade-level teachers to the Memphis iZone. In contrast, students in schools in both the top and bottom quartiles performed 0.30 to 0.31 standard deviations worse as a result of losing 100% of grade-level teachers to the Memphis iZone than those in schools in the middle half of socioeconomic status.

Table 4. Examining Moderating Effects by School Characteristics

Variable	By Percent ED			By Priority Status		
	Reading	Math	Science	Reading	Math	Science
Teacher Turnover to iZone	0.006	0.208	0.192***	-0.113*	-0.198	-0.088
	(0.063)	(0.378)	(0.035)	(0.049)	(0.140)	(0.067)
Teacher Turnover to iZone * Top 25% (Most Economically Disadvantaged)	-0.144	-0.347	-0.312***			
	(0.080)	(0.398)	(0.067)			
Teacher Turnover to iZone * Bottom 25% (Least Economically Disadvantaged)	0.012	-	-0.300***			
	(0.079)		(0.072)			
Teacher Turnover to iZone * Priority School				0.039	0.243	-0.010
				(0.086)	(0.235)	(0.112)
R squared	0.630	0.475	0.541	0.630	0.475	0.541
N	13,043	9,839	10,260	13,043	9,839	10,260

* p<0.05, ** p<0.01, *** p<0.001
Student Controls: Gender, Race, FRPL status, Special Education status, ELL status, Mobility Status, Prior Reading Test Score, Prior Math Test Score, Prior Science Test Score
School Controls: Percent Minority, Percent FRPL, Percent Student Mobility, School Level
All models include school-by-year fixed effects.

The remaining columns of table 4 examine if the effects of teacher turnover to iZone differentially affect lower-performing schools. In particular, we compare sending schools that have been labeled as part of the bottom 5% of schools in the state (Priority schools) to other sending schools. Priority schools generally have larger proportions of minority and economically disadvantaged students. Whereas non-Priority sending schools are 88% minority and 79% economically disadvantaged, Priority sending schools are 99% minority and 90%

economically disadvantaged. Further, whereas non-Priority sending schools score 0.32 to 0.50 standard deviations below the state average, Priority sending schools scored 0.84 to 1.12 standard deviations below average, depending on the subject. Therefore, based on previous literature, Priority sending schools should experience substantially larger negative effects than non-Priority sending schools.

The last three columns of table 4 show the results of interacting Priority school status with teacher turnover to the Memphis iZone. We find no differential effects for sending Priority schools relative to sending non-Priority schools.

Discussion

High-performing schools generally have a competitive advantage in the teacher labor market. These schools typically have students that are easier to educate, less accountability pressure, and better working conditions. Research shows that financial incentives have been a successful recruitment strategy for evening the playing field and making low-performing schools more competitive in attracting high-quality teachers. States and districts across the nation are relying on highly effective teachers to help turnaround their lowest-performing schools. In fact, two of the three federally-approved reform strategies previously discussed that allow schools to continue to operate require or, in practice, result in at least half of the teaching staff being replaced – turnaround and restart models. However, in a zero-sum teacher labor market, the number of highly effective teachers is fixed at least in the short term. Shuffling around these teachers results in positive impacts on the students who they now instruct, but can come at the expense of the students in the sending schools.

This study examined the systemic effect of teacher recruitment into the Memphis iZone on the students in sending schools. While there is some variation across models and power is an issue in some cases, the estimates are consistently negative but many are too imprecisely estimated to be statistically significant. Five out of nine estimates were statistically significant ranging from a -0.08 to -0.20 standard deviation change in student test scores as a result of 100% grade-level teacher turnover. Grades that lost reading teachers to the Memphis iZone lost 57% of their grade-level teachers, on average. In math and science, this percentage was 66% and 67%, respectively. Therefore, the average loss from this recruitment was approximately 0.05 to 0.13 standard deviations depending on model and subject. This leads us to conclude that any gains that iZone schools may be experiencing from the recruitment of high quality teachers is being partially offset by weaker performance in the sending schools and to some extent, creating a zero-sum game. In only one out of five specifications resulting in statistically significant findings was the effect of teacher turnover to the Memphis iZone greater in magnitude and statistically different from the effect of other teacher turnover in the same schools. This may suggest that the differences experienced in regard to severed working relationships, classroom disruptions, and changes in teacher quality were minimal between the two groups, or that the value-added effect as a result of these differences is small relative to the effect of losing a teacher regardless of effectiveness.

To further understand the impact of the iZone schools specifically and the practice of creating incentives of recruiting high quality teachers into turnaround schools more generally, it is important to examine the characteristics of the schools from which teachers were drawn, and whether these systemic effects vary across different school characteristics. If high quality teachers were pulled from other schools with large economically disadvantaged populations or low-performing schools, the unintended consequences of the teacher recruitment strategies for school turnaround could be more harmful than productive. Following previous literature, our analysis finds that schools in the top quartile of economically disadvantaged students (most economically disadvantaged) suffered greater losses than the middle half, particularly in science. However, the bottom quartile of economically disadvantaged students (least economically disadvantaged) also suffered comparable losses relative to the middle half, which runs contrary to previous research.

Particularly relevant to the Tennessee context are Priority schools, the state's lowest performing 5% of schools. 33% of sending schools were also Priority schools. If iZone schools were simply pulling the best teachers from other Priority schools, the school turnaround strategy can be counter-constructive if the sending Priority schools are performing even worse without these teachers. Our results, however, suggest that grades in the sending Priority schools are not performing any better or worse than those grades without the loss of these teachers to the iZone. This finding could be partially explained by regression to the mean – the performance of Priority schools may be so low that they cannot perform much worse. Nonetheless, it does not appear that the students in state's lowest-performing schools are adversely affected due to the loss of their highest quality teachers to the iZone.

It is unfortunate that some students have to lose in order for others to gain, but if only higher-performing schools are experiencing the negative consequence of teacher recruitment for low-performing schools, perhaps this strategy may be a good solution for advancing educational equity. To fully answer this question, further work should be done comparing the positive effects of the iZone schools to the negative effects on the sending schools, taking into account the types of students served. Future work should also investigate whether teacher turnover increased as a result of this recruitment program. Teachers who left for the Memphis iZone may have turned over without the incentives. Lastly, the long-term effects of teacher turnover should be examined. While students may experience a loss in the year directly after a teacher exits the school, schools may be able to recover over time by hiring an equally effective teacher or developing other teachers. Particularly for higher-performing schools, this recovery period may be rather short. If schools are able to rebound quickly, there may be even greater support for recruiting highly effective teachers from these environments to help turnaround low-performing schools.

References

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

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

Atteberry, A., Loeb, S., & Wyckoff, J. (2017). Teacher churning: reassignment rates and implications for student achievement. *Educational Evaluation and Policy Analysis*, *39*(1), 3-30.

Bifulco, R. (2012). Can Nonexperimental Estimates Replicate Estimates Based on Random Assignment in Evaluations of School Choice? A Within-Study Comparison. *Journal of Policy Analysis and Management*, *31*(3), 729-751.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2008a). *Who leaves? Teacher attrition and student achievement* (No. w14022). National Bureau of Economic Research.

Boyd, D., Lankford, H., Loeb, S., Ronfeldt, M., & Wyckoff, J. (2011). The role of teacher quality in retention and hiring: Using applications to transfer to uncover preferences of teachers and schools. *Journal of Policy Analysis and Management*, *30*(1), 88-110.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short careers of high-achieving teachers in schools with low-performing students. *The American economic review*, *95*(2), 166-171.

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2008b). The impact of assessment and accountability on teacher recruitment and retention: Are there unintended consequences?. *Public Finance Review*, *36*(1), 88-111.

Burnette II, D. (2017). A Passion for Saving Schools on the Brink. *Education Week*. Retrieved from https://leaders.edweek.org/profile/sharon-griffin-chief-of-schools-school-turnaround/

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. The American Economic Review, 104(9), 2633-2679.

Clotfelter, C. T., Glennie, E. J., Ladd, H. F., & Vigdor, J. L. (2008). Teacher bonuses and teacher retention in low-performing schools: Evidence from the North Carolina $1,800 teacher bonus program. Public Finance Review, 36(1), 63-87.

Cowan, J., & Goldhaber, D. (2015). Do bonuses affect teacher staffing and student achievement in high-poverty schools? Evidence from an Incentive for National Board Certified Teachers in Washington State. Center for Education Data & Research.

Darling-Hammond, L., & Sykes, G. (2003). Wanted, a national teacher supply policy for education: The right way to meet the" highly qualified teacher" challenge. *education policy analysis archives*, *11*, 33.

Dee, T. S., & Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from IMPACT. *Journal of Policy Analysis and Management*, *34*(2), 267-297.

Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdumy, S., Graczewski, C., ... & Giffin, J. (2017). School Improvement Grants: Implementation and Effectiveness. NCEE 2017-4013. *National Center for Education Evaluation and Regional Assistance*.

Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational psychologist*, *36*(2), 103-112.

Fowler, R. C. (2003). Massachusetts signing bonus program for new teachers. education policy analysis archives, 11, 13.

Glazerman, S., & Max, J. (2011). Do low-income students have equal access to the Highest-performing teachers? (No. b5869da6f9a54ef09fcd97be781b8a33). Mathematica Policy Research.

Glazerman, S., Protik, A., Teh, B. R., Bruch, J., & Max, J. (2013). Transfer Incentives for High-Performing Teachers: Final Results from a Multisite Randomized Experiment. NCEE 2014-4004. National Center for Education Evaluation and Regional Assistance.

Goldhaber, D., Gross, B., & Player, D. (2007). Are Public Schools Really Losing Their Best? Assessing the Career Transitions of Teachers and Their Implications for the Quality of the Teacher Workforce. Working Paper 12. National Center for Analysis of Longitudinal Data in Education Research.

Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. Educational Researcher, 44(5), 293-307.

Gonzales, J. (2016, February 17). Memphis iZone schools outpaced state-run turnaround efforts. *The Tennessean*. Retrieved from http://www.tennessean.com/story/news/education/2016/02/17/memphis-izone-schools-outpace-state-run-turnaround-efforts/78693182/

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin?. *Educational Researcher*, *39*(1), 59-68.

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

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. Journal of human resources, 39(2), 326-354.

Hanushek, E. A., Rivkin, S. G., & Schiman, J. C. (2016). Dynamic effects of teacher turnover on the quality of instruction. Economics of Education Review, 55, 132-148.

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

Henry, G. T., Fortner, C. K., & Bastian, K. C. (2012). The effects of experience and attrition for novice high-school science and mathematics teachers. *Science*, *335*(6072), 1118-1121.

Henry, G. & Redding, C. (2017). Explaining the Effects of Teacher Turnover on Student Achievement.

Henry G., Zimmer, R., Attridge, J., Kho, A. & Viano, S. (2015). Teacher and student migration in and out of Tennessee's Achievement School District. Retrieved from *https://peabody.vanderbilt.edu/research/tnedresearchalliance/files/ASD_Teacher_Student_Migration.pdf*

Henry, G., Zimmer, R., Kho, A. & Pham, L. (2017) Recruitment and Retention of Teachers in Tennessee's Achievement School District and iZone Schools. Retrieved from https://peabody.vanderbilt.edu/research/tnedresearchalliance/files/Teacher_Retention_in_ASD_and_iZone_vF.pdf

Isenberg, E., Max, J., Gleason, P., Potamites, L., Santillano, R., Hock, H., & Hansen, M. (2013). Access to Effective Teaching for Disadvantaged Students. NCEE 2014-4001. Institute of Education Sciences.

iZone: About Us. (2017). *Shelby County Schools*. Retrieved from http://www.scsk12.org/izone/

Jackson, C. K. (2012). Non-cognitive ability, test scores, and teacher quality: Evidence from 9th grade teachers in North Carolina (No. w18624). National Bureau of Economic Research.

Jackson, C. K., Rockoff, J. E., & Staiger, D. O. (2014). Teacher effects and teacher-related policies. Annu. Rev. Econ., 6(1), 801-825.

Kebede, L. F. (2016, May 2). iZone chief Sharon Griffin on fixing Memphis' most challenging schools. *Chalkbeat*. Retrieved from http://www.chalkbeat.org/posts/tn/2016/05/02/izone-chief-sharon-griffin-on-fixing-memphis-most-challenging-schools/

Koedel, C., & Betts, J. R. (2007). Re-examining the role of teacher quality in the educational production function (pp. 2007-03). National Center on Performance Incentives, Vanderbilt, Peabody College.

Liu, E., Johnson, S. M., & Peske, H. G. (2004). New teachers and the Massachusetts signing bonus: The limits of inducements. Educational Evaluation and Policy Analysis, 26(3), 217-236.

Nelson, J. R., Martella, R., & Galand, B. (1998). The effects of teaching school expectations and establishing a consistent consequence on formal office disciplinary actions. *Journal of Emotional and Behavioral disorders*, *6*(3), 153-161.

Oliver, R. M., & Reschly, D. J. (2007). Effective Classroom Management: Teacher Preparation and Professional Development. TQ Connection Issue Paper. *National Comprehensive Center for Teacher Quality*.

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

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

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

Sanders, W. L., Wright, S. P., & Horn, S. P. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. Journal of personnel evaluation in education, 11(1), 57-67.

Sass, T. R., Hannaway, J., Xu, Z., Figlio, D. N., & Feng, L. (2012). Value added of teachers in high-poverty schools and lower poverty schools. Journal of Urban Economics, 72(2), 104-122.

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.

Steele, J. L., Murnane, R. J., & Willett, J. B. (2010). Do financial incentives help low-performing schools attract and keep academically talented teachers? Evidence from California. Journal of Policy Analysis and Management, 29(3), 451-478.

Steele, J. L., Pepper, M. J., Springer, M. G., & Lockwood, J. R. (2015). The distribution and mobility of effective teachers: Evidence from a large, urban school district. Economics of Education Review, 48, 86-101.

Sullivan, K. (2013). One to Watch: Memphis iZone. *The New Teacher Project*. Retrieved from https://tntp.org/blog/post/one-to-watch-memphis-izone

Taketa, K. (2017). Teachers leave Hazelwood schools in large numbers after another controversial year. *St. Louis Post-Dispatch*. Retrieved from http://www.stltoday.com/news/local/education/teachers-leave-hazelwood-schools-in-large-numbers-after-another-controversial/article_bd1a3de9-7117-5138-b105-4fb03a69eb24.html

Tennessee Department of Education [TDOE]. (2013). Tennessee Department of Education to Grant Signing and Retention Bonuses to Highly Effective Teachers. Retrieved from http://www.tennessee.gov/news/34146.

Tillery, A. (2017, January 17). The iZone, How School Districts are Closing the Achievement Gap. *Ebony*. Retrieved from http://www.ebony.com/news-views/izone-schools-memphis-philadelphia#axzz4dEDmiuzp

United States Department of Education [USDOE]. (2010). 2010 SIG Final Requirements. Retrieved from https://www2.ed.gov/programs/sif/index.html

United States Department of Education [USDOE]. (2009). Race to the Top Program Executive Summary. Retrieved from https://www2.ed.gov/programs/racetothetop/executive-summary.pdf

United States Department of Education [USDOE]. (2016). Teacher Incentive Fund – Year 5 Grantee Profiles: Shelby Count Board of Education, Tennessee. Retrieved from https://www2.ed.gov/programs/teacherincentive/memphispro.pdf

Williams, R. (2016, May 17). Memphis iZone Beliebers? Not me…yet. *A Disgruntled Republican in Nashville.* Retrieved from http://www.adisgruntledrepublican.com/2016/05/ memphis-izone-belieber-not-meyet.html

Zimmer, R., Henry, G. & Kho, A. (2017) The Effects of School Turnaround in Tennessee's Achievement School District and Innovation Zones. *Educational Evaluation and Policy Analysis*, 0162373717705729.

CHAPTER 3

AN EVALUATION OF TENNESSEE'S CHARTER SCHOOLS: EXAMINING CHARTER SCOPE OF OPERATIONS

Introduction

Charter schools, as an alternative to traditional public schools (TPS), have grown tremendously in numbers over the past 25 years. The National Center for Education Statistics estimates over 6,800 charter schools serving over 2.8 million students throughout the United States in the 2015-16 school year (NCES, 2017). Despite their longevity and reach, the debate surrounding the effectiveness of charter schools continues. Positive effects of charter schools on student achievement have been found in Chicago, Arizona, Idaho, Florida, Wisconsin, Texas, New York City, Boston, and Indianapolis (Hoxby & Rockoff, 2004; Solmon & Goldschmidt, 2004; Ballou, Teasley & Zeidner, 2006; Sass, 2006; Witte et al., 2007; Booker et al., 2007; Hoxby, Murarka & Kang, 2009; Abdulkadiroglu et al., 2011; Nicotera et al., 2011). Negative effects have been found in Michigan, North Carolina, Texas, Milwaukee, and Los Angeles (Bettinger et al., 2005; Bifulco & Ladd, 2006; Carruthers, 2012; Hanushek et al., 2007; Zimmer et al., 2009; Lauen, Fuller & Dauter, 2014). And still a number of studies have found that charters perform on par with traditional public schools (Betts et al., 2006; Ballou, Teasley & Zeidner, 2006; Zimmer et al., 2009). The large variation in these studies makes it difficult to draw conclusions about the overall effectiveness of charter schools from extant evidence. Instead, Berends and colleagues (2008) state that researchers have been asking the wrong question and suggest redirecting our attention on charter schools. Rather than asking "Do charter schools work?", we should instead be asking, "Under what conditions do charter schools work?" Because charter schools and the students attending them vary, understanding what characteristics make effective charter schools and for whom they are effective is important to understanding how charter schools can create better opportunities for students. This study seeks to understand one distinguishing characteristic of charter schools – their scope of operations.

Charter schools can be established by national or local organizations. Those belonging to a national network, such as the Knowledge is Power Program (KIPP), Achievement First, Rocketship Education, and Uncommon Schools, have an established school model that has been well-developed and implemented in other parts of the nation which they mimic in new locations. In contrast, locally established charter schools are developed by and may be tailored to the needs of the local community. To inform the debate of national and local charter schools, I draw on theory from the community development literature, which contrasts need-based approaches with asset-based approaches. Kretzmann & McKnight (1996) posit that development that draws upon the assets of the community are more successful than solutions imposed by "outsiders." Based on this framework, local charter schools should be more effective at improving student outcomes than those opened by national CMOs.

Tennessee, the subject of this study, provides several interesting differences from the other states in which charters have been evaluated, conditions that would appear to be favorable to finding positive effects of charters. First, Tennessee was one of the most recent states to allow charter schools. Therefore, nationwide CMOs entering the state at this later stage came with established and tested school models. Second, according to a report by the School Choice Demonstration Project (Battdorff et al., 2014), Tennessee was the only state in which charter schools received more funding per pupil than traditional public schools in the same districts, in large part due to sizable private funding from philanthropic organizations. Lastly, Tennessee was ranked as one of the highest in charter school growth in 2015 (NAPCS, 2016). These characteristics make Tennessee an appealing setting for examining the heterogeneity of charter school effectiveness.

This study also goes a step further by answering the question, "For whom do these charter schools work?" and "For whom do these different types of charter schools work?" Though not originally created with the purpose of addressing achievement gaps, researchers have found that charter schools can be a solution for closing racial and income achievement gaps (Hoxby, Murarka & Kang, 2009; Bifulco & Ladd, 2006; Dobbie & Fryer, 2011). Therefore, it is important to understand for which populations charter schools are effective. In this study, I investigate their effects by school level, student race, income, special education status, and English language learner status.

Charter scope of operations (national vs. local) is a characteristic that can be identified easily through charter school applications. Therefore, in an era where charter school growth is ever increasing, the results of this research can provide new, valuable information for policymakers and charter authorizers. In the next section, I review the literature on charter school evaluation and draw upon a theory from the community development literature to inform how charter scope of operations can influence student outcomes. I then describe the Tennessee charter school landscape including a brief history of charter schools in the state and what they look like today. Next, I discuss the data and methods used for this study, which includes two rigorous methods that have been utilized throughout the charter school literature – a student fixed effects model and propensity score matching. I follow this section with results and a discussion of the findings.

Literature Review

Review of Charter School Studies

As previously discussed, the charter school evaluation literature is replete with evaluations of the effectiveness of charter schools. Table 1 provides a summary of a collection of these studies, each of which uses methods to identify a plausibly causal effect of charter schools. Because not all studies used the same student achievement units and effect sizes may have differed slightly by models, overall results are displayed as positive, negative, or null.

As seen in the table, the first generation of charter school evaluation studies has yielded mixed findings in answering the question “Do charter schools work?” In some cases, yes – positive results have been found in Boston, Idaho, Texas, Chicago, New York City, Milwaukee, Indianapolis, Florida, Arizona, and Wisconsin; in other cases, no – negative results have been found in Michigan, North Carolina, Texas, Los Angeles, and Chicago; and in still other cases, the effects are neither positive nor negative, though there are probably more null effects than identified here due to publication bias. In summary, the results vary greatly and few overall conclusions can be drawn regarding the charter sector as a whole.

A second generation of charter school literature can be more informative answering (1) “What are the characteristics of effective charter schools?” and (2) “For whom are these characteristics effective?” Charter scope of operations (national vs. local) is one defining characteristic of various charter schools that has yet to be rigorously examined in the literature. In the following sections, I highlight how national and local structures can differentially influence the effectiveness of charter schools. I then address the second question posed above, identifying previous literature on heterogeneous effects of charter schools by school level, student race and student poverty status.

Table 1. Summary of Prior Charter Evaluation Studies and the Effects of Charter School Attendance on Student Achievement

Author & Year of Publication	Location	Years of Data (Lagged Years)	School Levels	Method	Effect
*Abdulkadiroglu et al., 2011	Boston	2002-2008	M, H	Lottery	Positive
*Angrist et al., 2010	KIPP Schools in Lynn, MA	2006-2009	M	Lottery	Positive
*Angrist, Pathak & Walters, 2013	Massachusetts	2002-2011	M, H	Lottery	Positive for urban schools; negative for non-urban schools
Ballou et al., 2006	Idaho	2003-2005	E, M, H	Student FE	Positive for ES; Null else Charter schools do worse with greater years of experience
Berends et al., 2008	Anonymous urban school district	2003-2006	E, M, H	Propensity Score Matching; Random Effects	Negative in initial years of enrollment, but positive as years in Charter increases
Bettinger, 2005	Michigan	1997	E	Difference in Differences	Negative
Betts, 2006	San Diego	2002-2004	E, M, H	Student FE	Null overall Negative in initial years of ES enrollment, but positive as years in Charter increases
Bifulco & Ladd, 2006	North Carolina	1996-2002	E, M	Student FE	Negative
Booker et al., 2007	Texas	1995-2002	E, M, H	Student FE	Negative in first year, but positive in years 2 and 3
Carruthers, 2012	North Carolina	1996-2007	M	Student FE	Negative
CREDO, 2009	15 states + DC	Varies	Varies	Matching - Virtual Control Records	Overall Negative Varies by State
Cremata, et al, 2013	27 states + DC	Varies	Varies	Matching – Virtual Control Records	Overall Positive Varies by State
*Curto & Fryer, 2012	Washington, DC	2008-2009	M	Lottery	Positive
*Dobbie & Fryer, 2011	Harlem, New York City	2005-2007	E, M	Lottery	Positive
*Dobbie & Fryer, 2016	Texas	1995-2013	E, M, H	Matching and Fixed Effects	Null Overall Positive for No Excuses schools Negative for non-No Excuses schools

Eberts & Hollenbeck, 2001	Michigan	1997-1999	E	Fixed Effects	Negative
Gleason et al., 2010	15 states	2005-2008	M	Lottery	Null Overall Positive for Charter schools serving low-income or low-achieving students. Negative for Charter Schools serving high-income or high-achieving students.
Hanushek et al., 2007	Texas	1996-2002	M	Instrumental Variables	Negative in initial three years of opening Null in later years
*Hastings, Neilson & Zimmerman, 2012	Anonymous	2006-2009	E, M, H	Lottery	Positive
Hoxby & Rockoff, 2004	Chicago	2000-2002	E, M	Lottery	Positive
Hoxby, Murarka & Kang, 2009	New York City	2001-2008	E, M, H	Lottery	Positive
Imberman, 2011	Anonymous large urban district in the Southwest	1994-2005	E, M, H	Instrumental Variables	Negative
Lauen, Fuller & Dauter, 2014	Los Angeles	2003-2009	E	Multilevel Modeling	Negative
Lavertu & Witte, 2009	Milwaukee	2001-2007	E, M, H	Student FE	Positive for math in early years of charter
Miron et al., 2007	Delaware	2000-2005	E, M, H	Matching	Negative for ES Positive for MS/HS
Nicotera, Mendiburo & Berends, 2011	Indianapolis	2003-2006	E, M, H	Student FE	Positive
Sass, 2006	Florida	2000-2003	E, M, H	Student FE	Negative in initial years of operation Positive by Year 5
Solmon & Goldschmidt, 2004	Arizona	1998-2000	E, M, H	Multilevel Modeling	Positive Overall Positive for ES Null for MS Negative for HS
*Tuttle et al., 2013	KIPP Schools Nationwide	Varies	M	Propensity Score Matching; Lottery	Positive
Witte et al., 2007	Wisconsin	1998-2002	E, M	Student FE	Positive

Zimmer & Buddin, 2006	Los Angeles & San Diego	1998-2002	E, M, H	Student FE & Random-Growth Models	Negative for ES in San Diego Some Positive, Some Negative for MS/HS
Zimmer et al., 2009	Chicago, Denver, Milwaukee, Philadelphia, San Diego, Ohio, Texas	Varies	M, H	Student FE	Negative in Chicago & Texas Null everywhere else Charters do worst in first year of operation but improve over time

Charter Scope of Operations – National vs. Local

New charter schools are typically established through one of two means. One means includes an organization outside of a community, such as a nationwide charter management organization (CMO), who has created a well-developed charter model that has appeared to be successful in other settings. Looking to expand their success and reach more students, they open new schools in new locations. To implement their model with fidelity, they often bring along staff with experience implementing the model or staff who may be inexperienced but are wholeheartedly committed to the mission of the organization and are also new to the community. Some well-known national CMOs include KIPP, Achievement First, Rocketship Education, and Uncommon Schools.

Charter schools can also be opened by members of the local community. In some cases, these are independent charter schools that operate a single school. In other cases, the school may be part of a local CMO that has had experience operating schools in other neighborhoods. Regardless, the school model is tailored to the community and staff primarily come from the local community.

While national CMOs can provide a tested model, theory from the community development literature suggests that community-led reforms may be more effective. Traditional approaches to building communities utilize a "needs-based" approach (Kretzmann & McKnight, 1996). In a "needs-based" approach, the focus is on a need, deficiency, or problem in a community. It seeks to answer, "What is wrong with the community?" and "How can we fix it?" This approach, by design, associates the community with its negative qualities, defining its identity in this way. Outsiders view the community as being unable to solve their problems themselves and in *need* of assistance. As a result, these outsiders step in to "help", believing that they have or are the community's solution. While they may be completely sincere in their assistance, this outsider intrusion is perceived as patronizing and demoralizing by the community. It represents an outsider's assumption that they, the outsiders, know the community and the needs of the community more fully than the community itself and that outsiders do not believe the community has the capacity to address its own problems. In turn, the community may resist the outsiders' efforts, resulting in an outcome that runs counter to the outsiders' original goal of "fixing" the community.

The alternative approach to community development is an "asset-based" approach (Kretzmann & McKnight, 1996), where the emphasis is on the positive aspects of a community and the solutions are driven by the community. An "asset-based" approach seeks to answer, "What are the community's assets?" and "How can the community best utilize these assets to address problems within the community?" The approach does not mean that communities do not need additional outside resources, but that the resources will be more effective if the community itself is mobilized, invested, and engaged in the reform itself. It allows the community to determine what it needs most and empowers it to engage in the changes necessary. In turn, this process fosters pride in the community and ultimately the theory predicts that this leads to greater success in meeting the community's challenges.

The “needs-based” and “asset-based” approaches to community development can be applied in the education setting to the establishment of national CMO charter schools and local charter schools, respectively. National CMOs believe there is an unmet educational need in a community and that they can fill it. As a result, they, as outsiders, deliver their established charter model that has proved successful in other contexts, confident that the model will work in this new setting. Because of its prior success, they make few changes to the model, if any, to ensure that it can be replicated as closely as possible. Therefore, they take into account little of the community context and feedback from the community. At the same time, they bring and recruit new “outsider” staff whom they believe will make the charter school a success. Though their assistance may be completely altruistic, following Kretzmann & McKnight’s (1996) theory of a “needs-based” approach, their schools will result in little progress of improving student outcomes in the community. When there is progress, it may be unsustainable due to community resistance.

Locally grown and established charter schools, on the other hand, reflect more of an “asset-based” approach. These schools are formed by the community members and while some operate as a part of a CMO, the CMO only operates charter schools in that community. The people that staff these schools are themselves assets of the community, and the school or school model has been tailored for that specific community with the input of the community. Following Kretzmann & McKnight’s theory of community development, we might, therefore, expect local charter schools to be more effective at raising student achievement than national CMOs.

While this contrast of needs-based and asset-based approaches has not been previously cited in the school reform literature, the theory runs throughout common debates in education, such as state versus local control of schools. More recently, the theory was highlighted in a report on school turnaround (Glazer & Egan, 2016). Starting in 2012-13, Tennessee's state-run Achievement School District (ASD) began taking over low-performing schools in Memphis. The ASD turned many of these schools over to charter management organizations who would continue operating them as neighborhood schools. While the students mostly remained the same, nearly all of the teachers were replaced in these schools, about 2/3 by teachers new to the Memphis community (Henry et al., 2015; Zimmer, Henry & Kho, 2017). In evaluations of these state-takeovers, charter-conversion schools have yielded effects no better or worse than other low-performing schools in the state on student achievement measures. This is in contrast to consistently positive effects by a less intrusive district-run, "home grown" turnaround approach that did not involve state takeover of schools and only replaced a fraction of teachers, primarily from surrounding schools and school districts (Zimmer, Henry & Kho, 2017; Essay 2 – Systemic Effects). In a qualitative analysis examining reasons for this disparity, Glazer and Egan (2016) found that familiarity and engagement with the community and its history played a significant role in the success of the district-led initiatives and the lack of success of the outside organizations. The researchers summarized Memphis community members' critique of the state-led reforms as "the ASD lacks a nuanced feel for the unique culture and historical narratives of individual neighborhoods and that this has led to poor decisions and mistrust (2)." In describing these community members, the researchers said, "They resent the presumption that Memphis can 'be fixed' by outsiders (2)." The "outsider" approach was therefore a major obstacle to the success of the state-run reforms. Given the same Tennessee context, it may also be an obstacle for national CMOs that open new schools in the state.

1Heterogeneous Effects

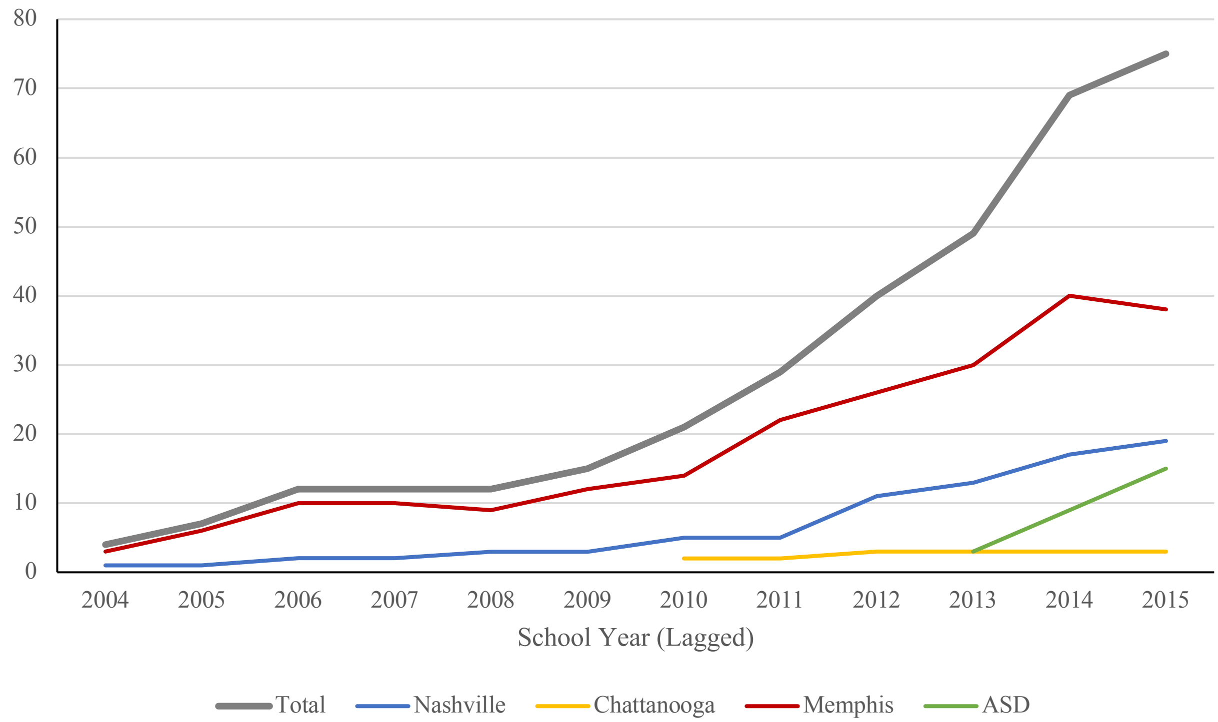
Researchers of charter schools argue that charter schools may help reduce achievement gaps for traditionally disadvantaged student populations (Hoxby, Murarka & Kang, 2009; Bifulco & Ladd, 2006; Dobbie & Fryer, 2011). Therefore, it is important to understand for which subgroups charter schools are effective as well as for which subgroups each charter school type is effective. Solmon and Goldschmidt (2004), Miron and colleagues (2007), and Zimmer and Buddin (2006) identify different effect estimates for different schools levels. In Arizona, charter schools serving elementary grades are performing better than traditional public schools. However, the charter middle schools are neither performing better nor worse, and the charter high schools are performing worse (Solmon & Goldschmidt, 2004). In Delaware, elementary charter schools are performing worse than traditional public schools while charter schools serving middle and high school grades are performing better (Miron et al., 2007). Zimmer & Buddin's work in California shows that Los Angeles' elementary charter schools are performing on par with traditional public elementary schools, but their secondary charters are performing higher in reading and lower in math. However, in San Diego, the elementary charter schools are performing worse, and the secondary charters are performing lower in reading and higher in math (Zimmer and Buddin, 2006). Following this literature, it is likely that charters, in general, may be more effective for certain subgroups of students, but the type of charter school and location may create differences.

A number of other studies have examined heterogeneous effects for various student subgroups. Gleason and colleagues (2010) found that charter schools serving lower-achieving and lower-income student populations perform better than traditional public schools, and charter schools serving higher-achieving and higher-income student populations perform worse than traditional public schools. CREDO's 2013 report also finds that charter schools have greater impacts for economically disadvantaged students, as well as black and Hispanic students, while their effect on white students are worse than traditional public schools. The same study also finds that charter schools do a better job educating English language learners and students with disabilities. On the other hand, California and Florida's charters targeting lower-achieving students do worse (Zimmer & Buddin, 2006; Sass, 2006), and those targeting special education students also perform worse (Sass, 2006). This variation in moderating effects may also be a function of the types of charter schools, which has yet to be addressed in the charter school research. Using data from the state of Tennessee, this study examines whether different types of charters better serves different subgroups of students.

Tennessee's Charter Schools

Tennessee joined the charter school movement in 2002, when it became the 40th state to pass legislation allowing charter schools into the state. The Tennessee Public Charter Schools Act of 2002 granted authorization rights to local education agencies. However, denied charter school applications could be appealed to the State Board of Education. In 2003, the first four charter schools in Tennessee opened – one in Nashville, three in Memphis. Over the following years, the number of charter schools increased gradually even though a few were closed. Figure 1 shows the number of charter schools over time in total and by location.

Figure 1. Number of Tennessee Charter Schools Over Time



As seen in Figure 1, charter schools initially began in Nashville and Memphis, with the number of schools generally increasing in both cities over time. In 2010, two charter schools opened in Chattanooga, the third city to receive charter schools. The last group indicated in Figure 1 is the Achievement School District (ASD), a state-run school district that took over low-performing schools and contracted school management out to CMOs. ASD charter schools are different from traditional charter schools in that they serve primarily as neighborhood schools rather than schools of choice and focus on school turnaround. Because of these differences, I exclude ASD charter schools from the study moving forward.

While it primarily appears that the number of charter schools continued to grow over the years, the dip from 2013-14 to 2014-15 for Memphis indicates that some charter schools closed. The first charter school closed was at the end of the 2006-07 school year in Memphis; the first in

Nashville was closed at the end of the 2009-10 school year. A second charter school closed in 2012-13 in Nashville, and seven closed in 2013-14 – two in Nashville, five in Memphis.

Table 2 provides some descriptive characteristics of Tennessee's charter schools. Of the 60 charter schools that were in operation in 2014-15, 18 charter schools serve elementary grades and 26 serve middle school grades only. Nine are high schools, and seven serve both middle and high school grades. All 60 operate in one of Tennessee's major cities – Memphis, Nashville, or Chattanooga. All but four charter schools are located in urban areas – two in rural areas and two in suburbs of these cities. Fifty of these schools serve student populations that are over 90% minority. Seven more fall between 70% and 89%. The last three serve 67%, 38%, and 28% minority student populations. There is greater variation in terms of percent of students economically disadvantaged. Across all cities, the average school percentage of economically disadvantaged students is about 75%, though the range extends from 32% to 98%. Over half of the charter schools in Nashville and Memphis serve student populations with 70% or more eligible for FRPM, but none of the charter schools in Chattanooga fall in this category.

Charters were either opened by a nationwide CMO or locally. I separate those that were opened locally into those belonging to a Tennessee-only network (local CMO) and those operating independently. As of the 2014-15 school year, about 13% of charter schools in the state were opened by a national CMO. In Nashville, the well-known KIPP operates 3 schools and Rocketship Education operates 1 school. KIPP operates 4 schools in Memphis. Six charter schools in Nashville are run by Nashville-specific operators – three by RePublic Schools and three by LEAD Public Schools. In Memphis, the majority of charter schools are operated by Memphis-specific networks – the W.E.B. DuBois Consortium of Charter Schools, City University Schools, Gestalt Community Schools, Promise Academies, Freedom Preparatory

Schools, Southern Avenue Charter Schools, the Memphis Business Academies, and the Memphis Academies of Health Sciences. Over 40% of charter schools in the state are operated independently without a specific network, including all three of Chattanooga's charter schools.

Table 2. Characteristics of Tennessee's 2014-15 Charter Schools

Characteristic	Nashville	Memphis	Chattanooga	Total
Elementary	5	12	1	18
Middle	12	14	0	26
Middle/High	1	5	1	7
High	1	7	1	9
Urban	18	36	2	56
Suburban	0	1	1	2
Rural	1	1	0	2
No. of Schools >70% Minority	17	38	2	57
Avg. School Percent Minority	85%	99%	72%	94%
No. of Schools >70% Eligible for FRPM	12	21	0	33
Avg. School Percent Eligible for FRPM	72%	77%	64%	75%
National CMO	4	4	0	8
Local CMO	6	20	0	26
Local Independent Charter	9	14	3	26
Total	**19**	**38**	**3**	**60**

In this study, I examine the overall and heterogeneous effects of charter schools in Tennessee. In particular, I investigate whether different charter scope of operations (national vs. local) differentially affects student achievement. Specifically, I seek to answer the following questions:

1) Compared to traditional public schools, how effective are Tennessee's charter schools at raising student achievement?

2) Which types of charter schools are more effective: national CMO schools or locally founded charter schools?
3) Does overall effectiveness and effectiveness by type of charter, national CMO-managed versus locally developed, vary by:
    a. School level (elementary, middle, high)?
    b. Student Race?
    c. Student Poverty status?
    d. Student Special education status?
    e. Student English language learner status?

Methods

Data & Measures

This study will utilize a student-level, longitudinal administrative dataset spanning from 2009-10 to 2014-15 provided by from the Tennessee Department of Education (TDOE) and managed by the Tennessee Education Research Alliance (TERA) at Vanderbilt University. The dataset includes demographic characteristics of students, standardized test scores, and school and grade assignment variables for all students in the Tennessee public school system for each school year.

The key dependent variables in this analysis are student test scores in reading, math, and science. In Tennessee, all students in grades three to eight are tested on an annual basis in reading, math, and science using the Tennessee Comprehensive Assessment Program (TCAP). High school students take end of course (EOC) exams upon completion of the course rather than a particular grade. Students receive scale scores for each of these exams. For this analysis, TCAP scale scores are standardized by year and grade; EOC scores are standardized by year, subject, and grade.

The key independent variable is whether a student attends a charter school. To determine if effects vary by national and local charter schools, I replace this variable with two indicator variables, one representing whether a student attends a national charter school and another representing whether a student attends a local charter school. Traditional public schools serve as the omitted category for both variables. In subsequent models, to assess differences between CMO-run schools and independent charter schools, I further split local charter schools into two indicator variables – local CMO-run schools and local independent charter schools. All national charter schools are CMO-run. To determine whether a charter school belongs to a network, I first begin with school names – many of the names identify the network, for example, KIPP Memphis Middle Academy. I also review school, district, and charter network webpages for those in which a charter network is not evident in the name.

Student characteristics are provided in TERA data and include gender, race, poverty status, special education status, and English language learner status. These are included as control variables and matching variables in the propensity score matching analysis.

School-level variables are aggregated up from the student-level to the school-level. The percent of students in poverty, the percent of students of minority status, school total enrollment are included as school-level control variables. Indicators of schools serving elementary, middle, and high school grades are also included to allow for estimating effects by school level.

Empirical Framework

Evaluating the effects of charter schools is a difficult task. Students cannot simultaneously be observed attending a charter school and the school they would have attended had the charter school not existed. Therefore, it is necessary to create a counterfactual for students who attend charter schools, which is easily recognized but extremely difficult to do convincingly in practice. Students who choose to attend charter schools are, by definition, different from those who do not choose to attend charter schools. They, or their families, made the choice to switch schools, which sets them apart from students who did not. As examples, these students may have more motivation, time, ability, and/or resources to do the research on which school to transfer to. Alternatively, they may have experienced a difficulty in their prior school and, as a result, chose to transfer to another school that would enroll them immediately, a charter school. Disentangling the effect of the charter school from these student and family characteristics that distinguish the students enrolling in charter schools from those who do not is difficult.

An ideal experiment for evaluating the effects of charter schools would randomly assign students to charter schools and others to traditional public schools. The difference in achievement between the treatment group, those that attend charter schools, and the comparison group, those who attend traditional public schools, would yield the impact of charter schools. However, this experiment is infeasible as it is not possible to assign students to schools randomly. As an alternative, as seen in Table 1, researchers have primarily used one or more of three quasi-experimental methods to identify a plausibly causal impact of charter schools – lotteries, student fixed effects, and/or matching. Each of these methods addresses the problem of selection bias by providing a properly balanced counterfactual in which the counterfactual would have the same outcomes as the treatment group if the treatment group was untreated or if the counterfactual group were treated. I discuss each of these strategies and their limitations below.

Lottery studies are one approach to creating a properly balanced counterfactual. These studies make use of lotteries that randomly select students to attend oversubscribed charter schools. Studies that use lotteries (Abdulkadiroglu et al., 2011; Angrist et al., 2010; Gleason et al., 2010; Hoxby & Rockoff, 2004; Hoxby, Murarka & Kang, 2009; Tuttle et al., 2013) pit lottery "winners," those who are randomly selected to have the opportunity to enroll in a charter school, against lottery "losers," those who are randomly denied the opportunity to enroll in the same charter school. Because both winners and losers applied for the lottery, researchers can make the assumption that winners and losers are similar on these motivational, situational, and/or familial characteristics that might normally threaten selection bias, making losers a balanced counterfactual for lottery winners, those who attend charter schools.

However, this methodology is only an option (1) for schools that are oversubscribed and (2) if schools maintain accurate records of their lotteries. In many cases, one or both of these cases cannot be met. Additionally, the literature has shown that charter schools that are oversubscribed are higher-performing charters – often that is the very reason they are oversubscribed (Abdulkadiroglu et al., 2011; Hoxby & Rockoff, 2004; Angrist, Pathak & Walters, 2013). Similarly, schools who maintain lottery records generally have high-quality record-keeping and organizational skills that may translate into or be indicative of high-quality instructional practices, particularly when records are not required by the district or state (Abdulkadiroglu et al., 2011). Therefore, the use of lotteries may affect the selection of charter schools into the study sample. In fact, reviewing the studies listed in Table 1 that utilize lottery methods to assess the effect of charter schools reveals that all of these studies yield positive effects of charter schools, with the exception of one by Gleason and colleagues (2010), which found null effects overall but positive effects for schools serving low-achieving and low-income students. Therefore, including only oversubscribed charter school in the study sample may yield results not appropriately generalized to the charter sector as a whole.

As an alternative, especially when lottery data are unavailable, as with this study, or the sample of oversubscribed schools are too limited, some researchers turn to student fixed-effect models for charter school evaluations. Student fixed-effect models allow for students to serve as their own comparison group, comparing students during periods they are enrolled in a charter school to periods when they are enrolled in the alternative setting, usually a traditional public school. This methodology, therefore, bypasses the problem of comparing a treatment group and comparison group that are not essentially equivalent. However, this approach has two weaknesses of its own. First, in order to be included in a student fixed-effect model, students must have had experiences in both the treatment and the alternative, in this case, a charter school and a traditional public school. Therefore, the effect is only identified for the "switchers". Students who always attend a charter school or never attend a charter school are excluded by this approach. This means that the outcomes for students in the study sample must be observed in both charter and traditional public schools. Therefore, in Tennessee where students begin taking standardized tests in the third grade, the study sample when using test scores as an outcome omits students who "switch" to or from a charter in grades 2 and below. These restrictions on the sample of students included in the study limits the generalizability of its results mainly to late elementary students and middle and high school students.

To determine if a sample of only switchers into or out of a charter school is representative of all charter school students including students who always attend a charter school, Zimmer and colleagues (2009) have compared the test score gains of students who switch between charter schools and traditional public schools to those who are always observed in a charter school, the former group being part of the fixed-effect estimates. In six of 14 location-by-subject comparisons, they find statistically significant differences between the two groups. The researchers conclude that switchers and students who always attend a charter school cannot be assumed to be the same and that their findings using the student fixed effect models cannot be generalized to all charter students.

Second, student fixed-effects models assume that students' pasts are good predictors of their futures. Researchers (Hoxby & Murarka, 2006; Zimmer et al., 2009) have argued that this may not be the case as students who transfer to or out of charter schools may have made the transfer in reaction to a particularly bad experience in the prior year. If so, this endogenous reaction could lead to selection into "treatment", biasing the estimates of the student fixed effects model. The bounce back from that experience would be incorrectly attributed to the school type they transferred to. To address this concern of selection bias, Zimmer and colleagues (2009) suggest testing for a systematic dip in the outcome the year prior to moving to a charter school. If found, one would be less confident in the results being attributable to the effect of charter schools.

The third alternative approach common in the charter school evaluation literature is matching. Propensity score matching can be used to create a comparison group of students that is similar to the treatment group based on observed covariates. To yield unbiased causal effects of charter schools, this method assumes that there exists a set of observable variables such that after matching a group of traditional school students on these variables, the outcomes are independent of selection into treatment. However, unobservable characteristics that may not be accounted for by observable characteristics, like motivation, could be correlated with both test score outcomes and choosing to attend a charter school. To the extent that these unobservable characteristics differentiate those students who enter charters and those who remain in traditional public schools, the effect estimates based on matching will be biased (in the case of motivation, biased upward). Several balance checks will be conducted to ensure that the treatment group and the comparison group are similar on observable characteristics, but it is not possible to assess whether those variables account for unobservable differences.

None of these strategies, on their own, yield an unbiased treatment effect estimate for all students who attend charter schools. Lottery estimates are only representative of high-performing schools; student fixed effect models exclude students who have always been enrolled in charter schools; and the estimates of propensity score matching can be biased by unobserved characteristics that are unaccounted for. However, to the extent that the effect estimates utilizing multiple strategies are consistent, we can be more confident in plausibly causal impacts. In this study, I utilize both student fixed-effect models and propensity score matching to identify the impact of charter schools on student achievement in Tennessee.

Student Fixed-Effect Models

The student fixed-effects approach is modeled as:

$$y_{igt} = \beta_0 + \beta_1 Charter_{ist} + S_{igt}\,\mathrm{B_j} + X_{st}\,\mathrm{B_k} + \eta_i + \theta_{gt} + e_{igst} \quad (1)$$

where y_{igt} represents the standardized test score for student i in grade g in year t. $Charter_{ist}$ is an indicator for whether student i in school s is enrolled in a charter school in year t. S_{igt} allows for time-varying student characteristics (economically disadvantaged, English language learner, special education, and mobility status) and X_{st} allows for school level characteristics (school percent economically disadvantaged, school percent minority, school total enrollment, and new school status). η_i captures individual student fixed-effects, and θ_{gt} captures year by grade fixed effects. The year by grade fixed effect is included to account for any differences in average growth in test scores across grades and years.

Students who switch schools, regardless of school type, may require some time to adjust as they get accustomed to the new environment. Following the work of Hanushek and colleagues (2007) and Zimmer and colleagues (2009), I include an indicator representing if student *i* is in his first year in school *s* in year *t*.

Additionally, prior research suggests that charter schools can be expected to perform worse in their first one to three years of operation (Betts, 2006; Sass, 2006; Booker et al., 2007; Hanushek et al., 2007; Berends et al., 2008). This makes sense as brand new schools, charter or not, may just be starting to put systems into place that will affect performance. To account for these early stages of a school, I include a dummy variable to indicate whether school *s* is in its first three years of operation in year *t* in a second analysis.

To determine if the effects of charter schools differ according to charter scope of operations, I replace $Charter_{ist}$ in equation (1) with two indicator variables – $LocalCharter_{ist}$ to represent if a student attends a local charter school and $NationalCharter_{ist}$ to represent if a student attends a national charter school. Traditional public schools serve as the omitted group.

Local charter schools can be operated as independent schools or part of a local CMO. Those that are a part of a local CMO belong to a network of schools, which makes it different from an independent charter school. Having multiple schools may be an indication of a successful relationship with the community, but it could also mean that the CMO may be replicating the needs-based approach of a national CMO but on a smaller scale from neighborhood to neighborhood. To determine if local CMO schools and local independent charter schools have different effects, I further disaggregate local charter schools into local independent charter schools and local CMO charter schools in subsequent analyses. In other words, I replace $Charter_{ist}$ in equation (1) with three mutually exclusive indicator variables - $LocalIndependentCharter_{ist}$, $LocalCMOCharter_{ist}$, and $NationalCharter_{ist}$. Note that all national charter schools, by definition, belong to a CMO. Traditional public schools serve as the omitted group.

When examining heterogeneous effects by school level and student characteristics, I include interaction terms between $Charter_{ist}$ and the respective moderating variables. For example, to determine if the impact of charter schools varies by economic disadvantage, I include $Charter_{ist}$, $EconDis_{igt}$, and $Charter_{ist} * EconDis_{igt}$. The coefficient on $EconDis_{igt}$ will provide the effect estimate of being economically disadvantaged on test scores for students while they are in a traditional public school, $Charter_{ist}$ will provide the effect estimate of being in a charter school for students that are not economically disadvantaged, and the interaction $Charter_{ist} * EconDis_{igt}$ will provide the effect estimate for economically disadvantaged students in charter schools, over and above the effect of being in an economically disadvantaged students.

To assess the validity of the student fixed effect approach and its assumptions, I conduct two tests previously discussed and suggested by Zimmer and colleagues (2009). To determine if the fixed effect sample of only switchers into or out of a charter school is representative of all charter school students including students who always attend a charter school, I will compare the mean test score gains of students who switch between charter schools and traditional public schools to those who are always observed in a charter school. To assess if the charter school effects are attributable to a shock in the year prior to moving to a charter school rather than the charter school itself, I falsely assume that students transfer to a charter school in the year prior and assess for an effect. Because students have in actuality not moved in this prior year, I should not find an effect. If an effect occurs in the prior year, it is not possible to attribute the effect of a charter school (in the correct year of transition) to the charter school itself.

The analytic sample for the fixed effect model consists only of Tennessee public school students who were in grades 3 and up at some point during the 2009-10 to 2014-15 school years. Because the outcome is test scores and students only test in these grades, we cannot include students in grades 2 and below. As previously discussed, the sample for the student fixed-effect model only includes students who have made “switches” between charter schools and traditional public schools.

Lastly, as noted earlier, several of the schools that started as charter schools in 2012-13 through 2014-15 came in as part of the ASD. These schools were different from traditional charters as they served primarily as neighborhood schools that were taking over low-performing schools. I exclude this group of schools from the analysis.

The student fixed-effect approach is limited in the generalizability of its estimates to only students that ever switch between charter schools and traditional public schools. I utilize a second procedure, propensity score matching (PSM) to help broaden the population of students to which I can extend these effects. To the extent that the two different procedures yield similar results, I can be more confident in their estimates as plausibly causal impacts of charter schools.

Propensity Score Matching

In this study, propensity score matching rests on the ability to match students in charter schools to students in traditional public schools based on observable characteristics that also removes bias that may be attributable to unobservable characteristics. There have been several studies that point to conditions under which the propensity score matching procedure produces results similar to randomized experiments and therefore appears to have removed unobserved confounders (Glazerman, Levy & Myers, 2003; Cook, Shadish & Wong, 2008; Shadish, Clark & Steiner, 2008; Bifulco, 2012). Glazerman, Levy & Myers (2003) assessed 12 within-study comparisons of experimental and non-randomized design estimates involving job training and employment service programs' effects on earnings to determine what characteristics of the studies yielded unbiased estimates of nonexperimental methods. The researchers found that including an extensive set of controls in a matching method helped to significantly reduce bias. In particular, prior earnings in these studies that used post-treatment earnings as the outcome of interest was the most important control for this purpose. In addition, matching based on geography in which the treatment and comparison groups come from similar geographical locations significantly helped to reduce bias. In subsequent studies, Cook, Shadish, and Wong (2008) and Shadish, Clark, and Steiner (2008) confirm that geographical matches, the use of pretest scores, and including an extensive set of covariates that are likely to predict entering treatment produce comparable results between randomized experiments and matching strategies. Lastly, using student-level data from magnet schools, Bifulco (2012) finds that including pretreatment test scores when comparison group students are drawn from the same districts or districts with similar student characteristics as those treated can reduce bias in nonexperimental methods between 64 and 96 percent. In sum, these studies strongly suggest that these three conditions, which I propose to implement in this study, can substantially reduce bias.

Based on these findings, I propose to implement a matching procedure that considers geographical location, pretreatment test scores, and other covariates that help predict likelihood of going into treatment and reaction to treatment. First, I restrict the pool of students eligible to be matched to students who are enrolled in charter schools to those who have never enrolled in a charter school that attend schools in the same districts. This will help ensure a geographical match. In particular, I exact match students on school district, as well as grade and year. Next, I estimate a logit regression model using observable student characteristics to generate a propensity score *p* for each student to enter a charter school. To generate propensity scores, I include the prior year's test score for the respective subject (reading, math, and science)[10] as well as gender, race, poverty status, special education status, English language learner status, and mobility status at the student level. At the school level, I include the percentage of minority students, the percentage of students that are economically disadvantaged, school total enrollment, and school level.

Next, charter students are matched with traditional public schools that minimize the difference in propensities to enroll in a charter school. Multiple matching designs have been used for reducing bias in treatment effect estimates in the literature. A study by Henry & Yi (2009) compares the effect estimates of fourteen different matching designs with estimates from the Project STAR random assignment study to determine the designs that were the most likely to produce estimates consistent with the experimental design estimates. Of the 14 procedures, ten were individual level matching procedures, including variations of nearest-neighbor matching (with and without calipers and replacement), one-to-five matching, radius matching, Gaussian and Epanechnikov kernel matching, and full matching. Only two of these were identified as adequate matching methods – a full matching design and a radius matching.

Following the work of Henry & Yi (2009) and recent literature in propensity score matching, I use radius matching with a radius of 0.001 to match charter school students with

[10] Rather than matching on all three subjects, I create three different matched samples for each of the three subjects. If I match on all three subjects, students would require prior test scores in all three, which would omit a large proportion of high school students given the testing procedures at the high school level. As described earlier, students test at the end of a subject rather than at the end of a grade. Therefore, a large proportion of students will not necessarily have prior year test scores in all three subjects in the same year.

traditional public school students to create the comparison group. Radius matching specifies a maximum propensity score difference between treatment observations and control observations. Differences that are larger than the specified difference, called the radius, are not included.

Following previous charter school evaluation literature that use propensity score matching techniques, I also use one-to-one nearest neighbor matching with replacement with a caliper of 0.01. While Henry & Yi (2009) did not find that the results of this procedure matched causal estimates, using the procedure provides comparability with previous charter school evaluation literature. Nonetheless, results in the following section do indicate similar findings regardless of either matching procedure.

After creating the matched samples, it will be important to check for balance across the treatment and comparison groups, which I do in several different ways. First, I compare the graphical distribution of propensity scores between the treatment and comparison groups. To ensure a balanced sample, the region of common support should overlap substantially and the distributions should look similar. If these two conditions are not met, there may be systematic differences between the treatment and comparison groups that bias the matching effect estimates. Following Rosenbaum & Rubin (1985), I also examine the standardized percentage bias between the matched treated and comparison groups for each matching variable. A well-matched sample should have a standardized percentage bias no greater than 5% for each variable. I can also observe the mean standardized percentage bias encapsulating all matching covariates as well. Particularly for continuous matching covariates, it is important to assess the variability of these variables across the treated and comparison groups. Following Rubin (2001), the variance ratio of the treated to comparison group should be ideally fall between 0.8 and 1.25. Rubin's B, which assesses the standardized difference of the means of the propensity score, and Rubin's R, which assess the variances of the propensity score, have also been used to assess for sufficiently balanced samples. Rubin (2001) recommends B to be less than 25 and R to be between 0.5 and 2 for a well-balanced sample. Lastly, the pseudo-R^2 of the logit estimation of treatment on the matching variables should be fairly small as the covariates should not be able to predict treatment after matching.

Assuming a well-balanced matched sample, I, lastly, run ordinary least-squares models including the treatment variable, propensity score and the matching covariates to determine effect estimates of charter schools:

$$y_{igt} = \beta_0 + \beta_1 \, Charter_{isdt} + S_{igt} \, \mathrm{B_j} + X_{sdt} \, \mathrm{B_k} + y_{igt-1} + p_{igt} + \delta_d + \gamma_g + \tau_t + e_{igst} \quad (2)$$

y_{igt} represents the standardized test score for student *i* in grade *g* in year *t*. $Charter_{isdt}$ is an indicator for whether student *i* in school *s* in district *d* is enrolled in a charter school in year *t*. S_{igt} represents a vector of student-level characteristics (gender, race, economically disadvantaged, English language learner, special education, and mobility status). X_{sdt} represents a vector of school-level characteristics (school percent economically disadvantaged, school percent minority, school total enrollment, school level, and new school status). y_{igt-1} represents the prior year test score for the subject being examined. p_{igt} represents the propensity for student *i* in grade *g* in year *t* to enter into a charter school. δ_d captures district fixed effects, γ_g captures grade fixed-effects, and τ_t captures year fixed effects.

Results

Student Fixed Effect Models

Research question 1 asks how effective Tennessee's charter schools are at raising student achievement relative to traditional public schools. The results of the student fixed effect models are displayed in table 3. Across all three subjects, students in charter schools experience greater achievement than students in traditional public schools. On average, students score about 0.15 standard deviations greater in reading, 0.16 standard deviations greater in math, and nearly 0.30 standard deviations greater in science relative to traditional public schools.

Table 3. Main Effects of Charter Schools – Student Fixed Effect Models

Variable	Reading	Math	Science
Charter	0.152***	0.156***	0.295***
	(0.023)	(0.040)	(0.042)
Economically Disadvantaged	0.007	-0.003	0.019
	(0.008)	(0.013)	(0.012)
English Language Learner	-0.180***	-0.139**	-0.061*
	(0.030)	(0.025)	(0.028)
Special Education	0.101***	0.048	-0.004
	(0.024)	(0.028)	(0.027)
Mobile	-0.024***	0.045***	-0.022*
	(0.007)	(0.009)	(0.009)
School Percent Economically Disadvantaged	-0.069	0.076	-0.074
	(0.046)	(0.103)	(0.112)
School Percent Minority	0.148*	0.261	0.080
	(0.072)	(0.141)	(0.137)
School Total Enrollment (in 100's)	0.003	0.011**	0.001
	(0.002)	(0.004)	(0.004)
Middle School	-0.065**	-0.066	0.098*
	(0.024)	(0.050)	(0.041)
High School	-0.150***	-0.109	-0.076
	(0.044)	(0.122)	(0.089)
Middle/High School Combination	-0.179**	-0.120	-0.073
	(0.055)	(0.098)	(0.082)
School Opened Within 3 Years	-0.054*	-0.032	-0.086
	(0.022)	(0.049)	(0.044)
Adjusted R squared	0.707	0.617	0.660
N	75090	69503	67301

Standard errors in parentheses, standard errors clustered at school level
* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Research question 3 asks how the effect of charter schools may vary for various school or student-level characteristics. The results of these moderating analyses are displayed in table 4. The first four rows are the results of examining effects by school level – elementary, middle, high, and combination middle/high schools. Because elementary schools serve as the omitted school level, “Charter School” indicates the estimated effect of charter elementary schools and effect estimates for interactions of “Charter * School Level” indicate the difference between charter elementary schools and charter schools at the respective school level (i.e., difference between charter elementary schools and charter middle schools). The results show that different school levels of charter schools do not appear to be significantly different in their performance.

The next five rows of table 4 show that charter schools are generally not any more or less effective for one race or another. There are two exceptions – black students in charter schools score 0.16 standard deviations lower in science than white students in charter schools, and students of other race in charter schools score 0.23 standard deviations lower in reading than white students in charter schools.

The last six rows of table 4 examine the effect of charter schools on particular disadvantaged groups of students – economically disadvantaged students, special education students, and English language learners. Charter schools do not appear to be more or less effective for economically disadvantaged students relative to non-economically disadvantaged students. For students with disabilities, charter schools appear to perform worse in science by 0.10 standard deviations, but no different in reading or math. However, charter schools are particularly effective for English language learners (ELL). ELLs in charter schools scored 0.24 standard deviations greater than their non-ELL charter school counterparts in reading. In math, they scored 0.26 standard deviations greater, and in science, they scored 0.12 standard deviations greater.

Table 4. Heterogeneous Effects of Charter Schools – Student Fixed Effect Models

Variable	Reading	Math	Science
Charter School	0.127	0.086	0.170
	(0.065)	(0.079)	(0.096)
Charter * Middle School	0.052	0.134	0.187
	(0.067)	(0.096)	(0.109)
Charter * High School	-0.012	-0.141	0.028
	(0.068)	(0.105)	(0.117)
Charter * Middle/High Combination School	-0.127	-0.068	-0.088
	(0.087)	(0.126)	(0.138)
Charter	0.204***	0.188	0.435***
	(0.052)	(0.099)	(0.064)
Charter * Black	-0.061	-0.045	-0.155**
	(0.047)	(0.088)	(0.053)
Charter * Hispanic	0.051	0.105	-0.025
	(0.049)	(0.085)	(0.052)
Charter * Asian	0.050	0.053	0.023
	(0.065)	(0.085)	(0.070)
Charter * Other Race	-0.226*	-0.045	0.218
	(0.115)	(0.157)	(0.163)
Charter	0.149***	0.150**	0.279***
	(0.026)	(0.045)	(0.047)
Charter * Economically Disadvantaged	0.003	0.006	0.018
	(0.016)	(0.027)	(0.023)
Charter	0.156***	0.159***	0.301***
	(0.022)	(0.040)	(0.042)
Charter * Special Education	-0.052	-0.051	-0.099***
	(0.028)	(0.030)	(0.029)
Charter	0.148***	0.150***	0.292***
	(0.023)	(0.040)	(0.042)
Charter * English Language Learner	0.239***	0.256***	0.120**
	(0.032)	(0.042)	(0.044)
N	75090	69503	67301

Standard errors in parentheses, standard errors clustered at school level
* p<0.05, ** p<0.01, *** p<0.001

Community development theory posits that charter schools that are founded locally and by the local community can be more effective than those managed by an outsider, national CMO approach. Table 5 displays the results of disaggregating charter schools into these two groups. The first two rows show that both locally- and national charter schools are more effective than traditional public schools. Locally-founded charter schools score 0.14 standard deviations greater than traditional public schools in reading and math. In science, locally-founded charter schools score 0.28 standard deviations greater than traditional public schools. National charter schools score 0.23, 0.28, and 0.42 standard deviations greater than traditional public schools in reading, math, and science, respectively. To answer research question 2, F-tests were used to test differences between locally-founded charter schools and national charter schools. Statistically significant differences are indicated in bold. While the effect estimates for national charter schools appear to be larger than locally-founded charter schools, in no subjects were significant differences found.

It is possible that being a part of a charter management organization may affect one's effectiveness, so I further disaggregate charter schools into three groups – independent charters that are locally operated, charter schools run by local CMOs, and charter schools run by national CMOs. These results are displayed in the next three rows of table 5. All three charter types yield effects greater than traditional public schools except independent charter schools in math. Independent charter schools score 0.10 standard deviations greater in reading and 0.23 standard deviations greater in science relative to traditional public schools. Local CMO charter schools score 0.16, 0.18, and 0.30 standard deviations greater than traditional public schools in reading, math, and science, respectively. Results for national charter schools are comparable to the analyses above. F-tests between each combination of charter school types reveal that national CMO charter schools outperform independent charter schools by 0.13 standard deviations in reading and 0.23 standard deviations in math, both of which are statistically significant. Local CMO charter schools outperform independent charter schools by 0.14 standard deviations in math. All three charter schools perform comparably in science.

Table 5. Effects by Scope of Operations – Student Fixed Effect Models

Variable	Reading	Math	Science
Locally-Founded Charter School	0.143***	0.140***	0.278***
	(0.022)	(0.038)	(0.040)
National Charter School	0.228***	0.279**	0.415***
	(0.051)	(0.084)	(0.081)
Independent Charter	0.096*	0.038	0.234***
	(0.037)	(0.062)	(0.064)
Local CMO	0.161***	**0.179*****	0.295***
	(0.024)	**(0.042)**	(0.046)
National CMO	**0.224*****	**0.270****	0.412***
	(0.052)	**(0.085)**	(0.081)
N	75090	69503	67301

Standard errors in parentheses, standard errors clustered at school level
* p<0.05, ** p<0.01, *** p<0.001
Bolded coefficients indicate statistically significant difference relative to independent charter schools.

As previously discussed, the student fixed effect models require two additional validity checks. First, because only "switchers" are included for these models, it is unclear whether the effect estimates are only generalizable to students who switch between traditional public schools and charter schools or whether they can also speak to students who have always attended charter schools. To assess its external validity, I model the effect of always attending a charter relative to switching to a charter school. In reading and science, I find insignificant differences of 0.004 and 0.009 standard deviations, respectively. In math, I find a significant difference with students always attending charter schools performing 0.03 standard deviations higher. This significant difference may indicate some difference between the two groups, but is fairly small in magnitude. This indicates confidence in the results of the above student fixed effect analyses to have broader generalizability beyond switchers and to include all charter school students.

Second, the gains for students attending charter schools could be the result of recovering from a bad experience in the year prior to moving to the charter school that may have spurred the transition to a charter school, rather than the effect of the charter school itself. To assess this, I falsely assume the move occurred in the year prior to students actually moving. In this case, I find no effects in the year prior with estimates all less than 0.01 and statistically insignificant. This helps to attribute the effects of charter schools found previously to the charter schools themselves.

Propensity Score Matching

An alternative approach to student fixed effect models for charter school evaluations uses propensity score matching. For each of three subjects (reading, math, and science) and two matching procedures (radius and nearest neighbor matching), I create a matched sample, yielding six different matched samples. Before conducting any analyses, it is important to first check for balance in these samples. In table 6, I display the average characteristics of the matched sample using math test scores and radius matching. As can be seen, the propensity score matching procedure yielded a treated and comparison group that are similar on the observed characteristics, with mean differences no greater than one percentage point. The matched sample meets sufficient balance according to the various tests of balance – the standardized percent bias is lower than 5% for all variables with the overall average percent bias at 1.2%, the variance ratios of continuous variables fall within the range of 0.8 to 1.25, Rubin's B falls well-below 25 at 7.1, Rubin's R falls within the range of 0.5 and 2, and the pseudo-R^2 of the estimation of treatment on matching covariates is sufficiently small at 0.001.

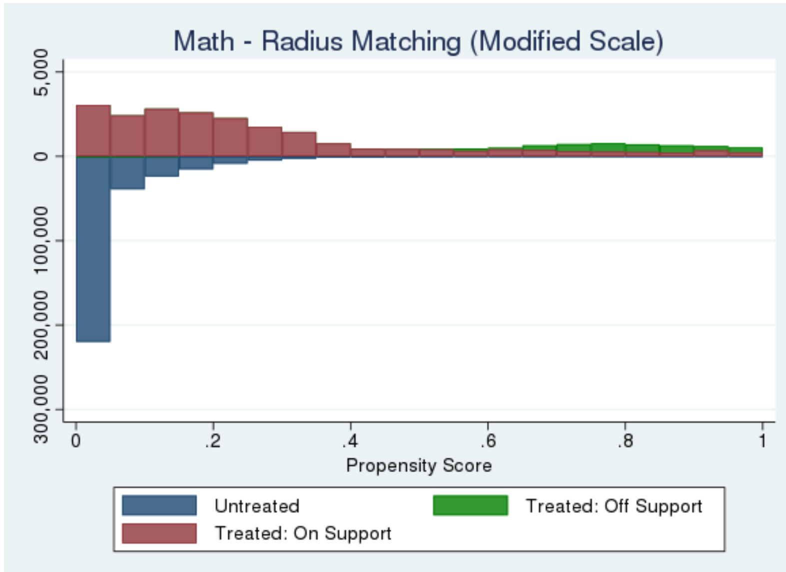
Table 6. Checking Balance for Math Radius Matched Sample

Characteristics	Mean		Mean Difference	Percent Bias	Variance Ratio
	Treated	Control			
Male	0.475	0.483	-0.009	-1.7	
Black	0.831	0.825	0.007	1.6	
Hispanic	0.106	0.108	-0.002	-0.6	
Asian	0.007	0.008	0.000	-0.3	
Other Race	0.000	0.001	0.000	-0.7	
Economically Disadvantaged	0.861	0.853	0.008	1.9	
English Language Learner	0.052	0.054	-0.001	-0.7	
Special Education	0.054	0.059	-0.004	-1.7	
Prior Mobile Status	0.195	0.195	0.000	-0.1	
Prior Math Score	-0.412	-0.406	-0.006	-0.6	0.94*
Prior School Percent ED	0.847	0.839	0.008	3.8	0.93*
Prior School Percent Minority	0.914	0.911	0.003	1.3	0.96*
Prior School Total Enrollment (in 100's)	4.765	4.799	-0.034	-1	1.07*
Prior School Middle School	0.433	0.431	0.002	0.3	
Prior School High School	0.022	0.022	0.000	0.1	
Prior School Middle/High School Combo	0.048	0.049	-0.001	-0.5	
Overall Measures					
Mean Percent Bias	1.2			**Rubin's B**	7.1
Pseudo-R squared	0.001			**Rubin's R**	0.82

One additional visual check for balance examines the distribution of propensity scores for the treated and comparison groups. Figure 2 provides this illustration. Note that the y-axis is proportional by group in order to properly assess the distribution within each group, and thus the treated and untreated distributions are not on the same scale relative to one another.[11] Given that less than 5% of Tennessee's students attend a charter school, the distribution for the treated and comparison groups look similarly distributed and share a sufficient region of common support.

[11] Because only 5% of Tennessee's students attend a charter school, it is difficult to assess the distributions of treated and untreated groups when using proportional scales. Nonetheless, I have included the figures with proportional scales in the Appendix.

Figure 2. Common Support for Math Radius Matched Sample



Note: The y-axis is proportional by group in order to properly assess the distribution within each group. The treated and untreated are not on the same scale.

Together, these various balance checks provide evidence that the treated and comparison groups are sufficiently balanced on observable characteristics. Similar tables and figures are available upon request for the two other subjects with radius matching and for all three subjects with nearest neighbor matching procedures. All matched samples appear to be sufficiently balanced according to the indicated checks.

Given sufficient balance, the research questions can be examined using the created samples. Table 7 summarizes tables 3 and 5 except that it displays the results of the propensity score matching models rather than the student fixed effect models. Row 1 shows that regardless of the matching procedure and subject, students in charter schools perform better than students in traditional public schools who are similar on observable characteristics. In reading, charter school students have gains of about 0.14 standard deviations greater than TPS students; in math, about 0.18 to 0.19 standard deviations; in science, about 0.23 to 0.26 standard deviations.

Table 7. Main Effects of Charter Schools and Effects by Scope of Operations – Propensity Score Matching

Variable	Reading		Math		Science	
	Radius Matching	NN Matching	Radius Matching	NN Matching	Radius Matching	NN Matching
Charter	0.144***	0.136***	0.191***	0.176***	0.263***	0.227***
	(0.025)	(0.026)	(0.044)	(0.045)	(0.048)	(0.046)
Locally-Founded Charter School	0.141***	0.133***	0.187***	0.172***	0.259***	0.222***
	(0.023)	(0.023)	(0.043)	(0.042)	(0.048)	(0.045)
National Charter School	0.170*	0.162	0.230*	0.214	0.309*	0.271*
	(0.085)	(0.085)	(0.107)	(0.118)	(0.127)	(0.117)
Independent Charter	0.097	0.100*	0.095	0.089	0.250**	0.206*
	(0.050)	(0.044)	(0.091)	(0.088)	(0.093)	(0.081)
Local CMO	0.150***	0.141***	0.206***	0.193***	0.260***	0.226***
	(0.026)	(0.025)	(0.046)	(0.045)	(0.052)	(0.051)
National CMO	0.164	0.159	0.216*	0.204	0.308*	0.269*
	(0.085)	(0.085)	(0.109)	(0.118)	(0.126)	(0.116)
N	316313	53093	267846	47470	248786	46010

Standard errors in parentheses, standard errors clustered at school level; NN = Nearest Neighbor
* p<0.05, ** p<0.01, *** p<0.001
Differences between locally-founded charter schools and national charter schools were not statistically significant. Differences between any combination of independent charter schools, local CMO charter schools and national CMO charter schools were not statistically significant.

Rows 2 and 3 of table 7 indicate that both locally-founded and national charter schools are more effective than traditional public schools – between 0.13 and 0.26 standard deviations for locally-founded charter schools, depending on subject, and between 0.17 and 0.31 standard deviations for national charter schools, depending on subject. While national charter schools may have larger effect estimates, the estimates are not statistically significant in two of six cases. F-tests yield no significant differences between locally-founded charter schools and national charter schools.

When disaggregating locally-founded charter schools into independent and local CMO charter schools (rows 4 through 6 of table 7), the results of F-tests again do not show any statistically significant differences between each combination of charter school types. Comparing to TPSs, however, all three are more effective in various models. In particular, local CMOs consistently show improved student achievement across subjects and models – 0.14 to 0.15 standard deviations greater than TPS in reading, 0.19 to 0.21 in math, and 0.23 to 0.26 in science. While the effects for national CMO charter schools are in most cases larger than local CMOs, the smaller sample size for this group leads to imprecise estimates in some cases. For independent charter schools, the effects are smaller and only sometimes significant.

Table 8 answers research question 3 and is comparable to table 4 except that it uses propensity score matching models rather than student fixed effect models. The table shows that, in reading and science, charter middle schools appear to be a heavy driver in the overall effect of charter schools found in table 7. Charter middle schools are 0.15 to 0.19 standard deviations more effective in reading and 0.30 to 0.32 standard deviations more effective in science than elementary charter schools.

Black students in charter schools perform significantly worse than white students in charter schools in science – a difference of 0.21 to 0.26 standard deviations. However, an F-test (not shown here) shows that black students in charter schools perform significantly better than black students in traditional public schools. They perform comparably to white students in reading and math. With few exceptions, charter school students of Hispanic, Asian, or another race also generally perform comparably with white students in charter schools.

According to the propensity score matching models, each of the last three subgroups analyzed – economically disadvantaged students, students with disabilities, English language learners – appear to benefit somewhat more than their respective counterparts from attending charter schools, depending on subject and matching method. Economically disadvantaged students attending charter schools score 0.05 standard deviations greater in reading, about 0.10

standard deviations greater in math, and 0.08 standard deviations greater in science than non-economically disadvantaged students in charter schools (though this is not consistently significant in reading and science). Special education students attending charter schools score 0.12 to 0.18 standard deviations greater in reading and about 0.14 standard deviations greater in science than regular education students in charter schools (not significant in math or in one model of science). English language learners score 0.11 to 0.13 standard deviations greater in reading than non-ELL students in charter schools.

Table 8. Heterogeneous Effects of Charter Schools – Propensity Score Matching

Variable	**Reading**		**Math**		**Science**	
	Radius Matching	NN Matching	Radius Matching	NN Matching	Radius Matching	NN Matching
Charter School	-0.015	0.014	0.043	0.024	0.003	-0.011
	(0.049)	(0.048)	(0.104)	(0.096)	(0.076)	(0.068)
Charter * Middle School	0.194***	0.153**	0.198	0.205	0.317***	0.299***
	(0.058)	(0.058)	(0.114)	(0.107)	(0.093)	(0.085)
Charter * High School	0.128*	0.087	-0.094	-0.066	0.172	0.106
	(0.056)	(0.058)	(0.132)	(0.125)	(0.110)	(0.108)
Charter * Middle/High Combination School	0.118	0.068	0.137	0.162	0.121	0.075
	(0.081)	(0.075)	(0.155)	(0.152)	(0.127)	(0.119)
Charter	0.180***	0.169***	0.276***	0.246**	0.486***	0.408***
	(0.042)	(0.043)	(0.077)	(0.079)	(0.084)	(0.080)
Charter * Black	-0.049	-0.043	-0.106	-0.089	-0.263***	-0.214**
	(0.039)	(0.040)	(0.071)	(0.074)	(0.076)	(0.076)
Charter * Hispanic	0.044	0.041	0.039	0.060	-0.031	-0.010
	(0.041)	(0.038)	(0.077)	(0.076)	(0.057)	(0.062)
Charter * Asian	0.003	-0.040	-0.194*	-0.168	-0.154	-0.041
	(0.075)	(0.090)	(0.079)	(0.104)	(0.081)	(0.100)
Charter * Other Race	0.030	-0.342	-0.149	0.509	0.012	0.014
	(0.275)	(0.405)	(0.413)	(0.434)	(0.183)	(0.167)
Charter	0.098**	0.106**	0.099*	0.090	0.189**	0.168**
	(0.031)	(0.037)	(0.049)	(0.050)	(0.058)	(0.060)
Charter * Economically Disadvantaged	0.052*	0.034	0.102***	0.098**	0.084*	0.067
	(0.022)	(0.030)	(0.028)	(0.031)	(0.037)	(0.041)
Charter	0.137***	0.126***	0.190***	0.174***	0.259***	0.219***
	(0.025)	(0.025)	(0.044)	(0.045)	(0.048)	(0.046)
Charter * Special Education	0.119**	0.177**	0.016	0.043	0.080	0.143*
	(0.041)	(0.062)	(0.041)	(0.057)	(0.044)	(0.067)
Charter	0.140***	0.132***	0.186***	0.172***	0.258***	0.223***

	(0.025)	(0.025)	(0.044)	(0.045)	(0.048)	(0.046)
Charter * English Language Learner	0.112*	0.133**	0.120	0.109	0.119	0.101
	(0.049)	(0.051)	(0.069)	(0.073)	(0.068)	(0.066)
N	316313	53093	267846	47470	248786	46010

Standard errors in parentheses, standard errors clustered at school level; NN = Nearest Neighbor
* p<0.05, ** p<0.01, *** p<0.001

Heterogeneous Effects Across Charter Scope of Operations

Lastly, I examine whether independent, local CMO, or national CMO charter schools are differentially effective for each of the moderating characteristics discussed previously. To conserve space, I do not include the full results here, but will highlight findings that are consistent across both the student fixed effect and propensity score matching models:

(1) Independent charter high schools appear to be less effective than other independent charter schools. They do particularly poorly in math, scoring up to 0.90 standard deviations lower than other independent charter schools.

(2) Local CMO middle schools are particularly effective, scoring up to 0.20 standard deviations greater in reading and 0.30 standard deviations greater in science than other local CMO charter schools.

(3) The positive effects of national CMO schools are driven primarily by elementary and middle schools. National CMO high schools generally perform on par with traditional public high schools.

(4) Independent charter schools and national CMO charter schools are particularly effective for ELL students with gains up to 0.30 standard deviations in reading for both charter school types.

(5) There are no consistent, significant differences in student achievement for students of different races, socioeconomic status, or special education status by charter scope of operations.

Discussion

Charter schools have been a controversial education reform strategy since they first originated. The first generation of evaluation work has yielded mixed findings of charter schools across different settings. This makes sense as charter schools were created to innovate educational approaches, and different approaches should yield different results. To better understand this variation, we should focus on which types of charter schools are effective and for whom they are effective. This study attempts to do that using data from Tennessee, one of the latest states to adopt charter schools. Because of this late adoption, Tennessee has had the advantage of having allowed other states and various charter management organizations to refine their strategies before entering Tennessee.

Table 9 summarizes the results of this study. Overall, I find that charter schools are more effective at improving student achievement than traditional public schools in the same districts the charter schools operate. When examining heterogeneous effects by school level, charter middle schools, if anything, have been particularly effective, though the results are inconsistent across models. Students of minority races generally benefit comparably to their white student counterparts from the charter school experience with one outstanding exception – black students do not perform as well as white students in science. Some inconsistencies lie among the results for economically disadvantaged and special education students. They generally perform either better or on par with their counterparts. English language learners, however, gain more from the charter experience than non-ELL students. These results indicate that charter schools are effective overall for students and particularly ELL students. In addition, they may be closing the income achievement gap, but do not appear to be closing the racial achievement gap.

Table 9. Summary of Findings

	Reading	Math	Science
Overall Charter Effect	+ +	+ +	+ + +
By Scope of Operations			
Independent Charter	+ / 0 / +	0	+ + +
Local CMO Charter	+ +	+ +	+ + +
National CMO Charter	+ + + / 0 / 0	+ + + / + + + / 0	+ + +
By School Level (relative to Elementary School)			
Middle School	0 / + + / + +	0	0 / + + + / + + +
High School	0 / + + / 0	0	0
Middle/High Combination	0	0	0
By Student Race (relative to White students)			
Black	0	0	– – / – – – / – – –
Hispanic	0	0	0
Asian	0	0 / – – / 0	0
Other Race	– – / 0 / 0	0	0
Economically Disadvantaged	0 / + / 0	0 / + + / +	0 / + / 0
Special Education	0 / + + / + +	0	– / 0 / + +
English Language Learner	+ + + / + + / + +	+ + + / 0 / 0	+ + / 0 / 0

(+) Statistically significant effect size > 0 and ≤ 0.10 standard deviations
(+ +) Statistically significant effect size > 0.10 and ≤ 0.20 standard deviations
(+ + +) Statistically significant effect size > 0.20 standard deviations
(–) Statistically significant effect size ≥ -0.10 and < 0 standard deviations
(– –) Statistically significant effect size ≥ -0.20 and < -0.10 standard deviations
(– – –) Statistically significant effect size < -0.20 standard deviations
0 Statistically insignificant
Shaded cells indicate consistent results across student fixed effect and both propensity score matching models. Unshaded cells indicate inconsistent results with student fixed effect results displayed first followed by radius matching results followed by nearest neighbor matching results.

In this study, I also utilize theory from the community development literature to inform the investigation of the role of community in charter school effectiveness. Assets-based community development suggests that charter schools that are founded locally are more effective than those founded nationally because they have greater context of the community in which the school operates and utilize more talent and resources from the community. However, the results of this study run contrary to this theory. While both types of charters are more effective than traditional public schools, there were no significant differences between the performance of national charter schools and locally-founded charter schools. One of the possible reasons for this result could be that national charter schools have a charter management organization to rely on that can provide assistance and support. To address this, I separated locally-founded charter schools into independent charter schools and charter schools belonging to local CMOs and found that the network factor may explain a proportion of the effectiveness as local CMO charter schools appeared, if anything, to perform better than independent charter schools. However, national CMO charter schools outperformed local independent charter schools. The success of national CMO schools in Tennessee could be attributed to the consistent approach they use. While their school models may not be specifically tailored to the communities in which they operate, these national models have been proven effective in other settings that may be similar in terms of the types of students and families they serve (i.e., high poverty schools serving underrepresented minorities). Thus, *geography* may not matter if their models effectively serve similar contexts.

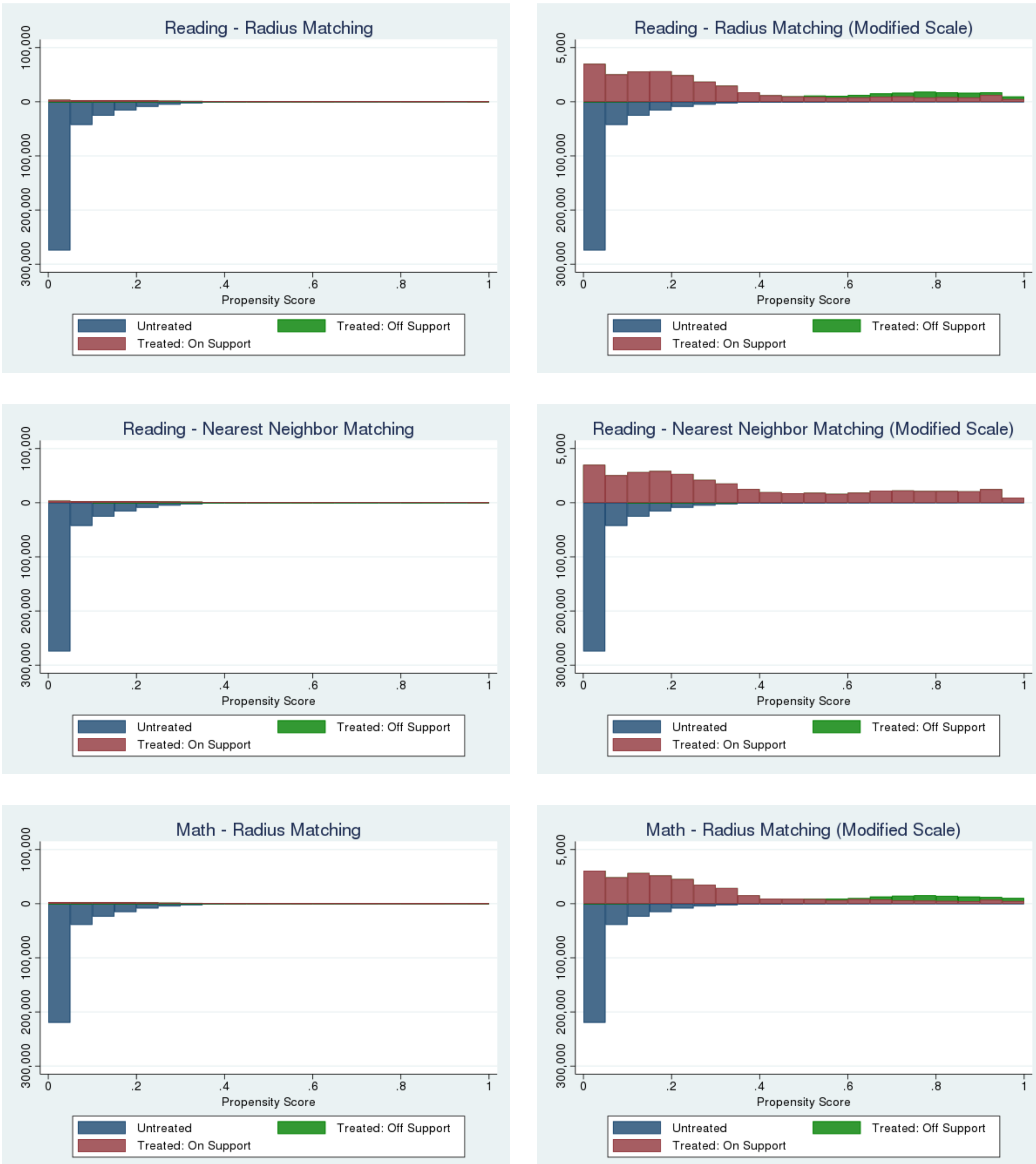
Another plausible explanation for the difference in hypothesized results could be that while local charter schools are founded and operated by the community, staffing for the schools hire new, inexperienced teachers. Thus, the lack of experience implementing the school vision may run counter to the advantage of the community-based leadership and vision. Further work should assess the staffing of different various charter school types to see if implementation is in line with the respective needs-based or assets-based approach.

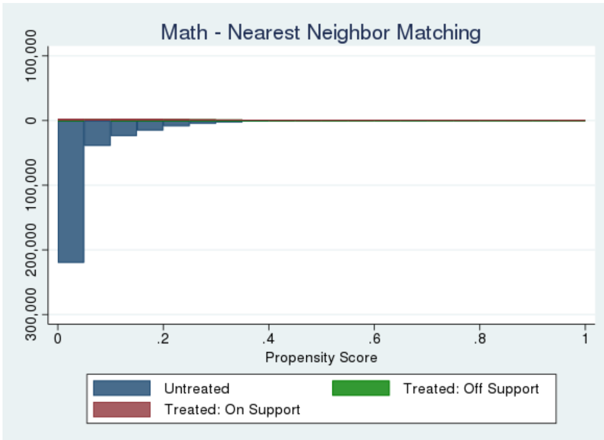
Conclusions about heterogeneous effects of the different types of charter schools were fairly consistent across the propensity score matching and student fixed effect approaches. For elementary schools, national CMO charter schools appeared to be the most effective. For middle schools, local and national CMO charter schools appeared to be the most effective. All charter types were effective for ELL students but independent charter schools and national CMO charter schools were particularly effective for this student subgroup.

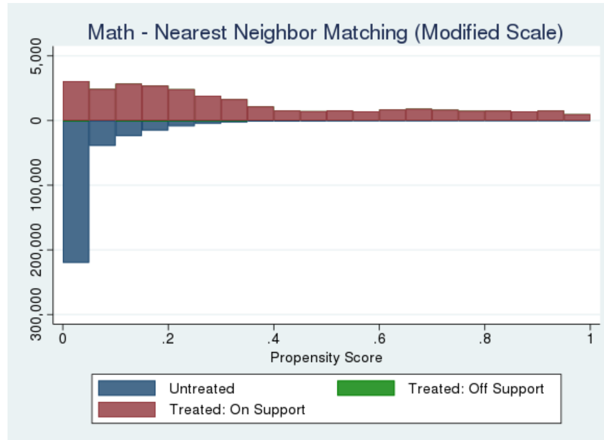
The results of this study are salient for charter authorizers going forward particularly in Tennessee. While all charter schools were, on average, more effective than traditional public schools, some were more effective than others. For instance, if looking to open a new middle school, local CMOs may have an advantage. The fact that many of the local middle schools are operated by the same CMO may serve as an advantage in regards to their experience with the particular school level. If a new charter application seeks particularly to address English language learners, perhaps independent charter schools or a national CMO charter school should receive more attention. This study has addressed this one major differentiator between charter schools and has also assessed the effects of charter schools across various student subgroups. Future research should address other characteristics of charter schools to help differentiate promise in charter schools. For instance, similar to the local versus national scope of operation and whether the charter school belongs to a CMO-network, there are other characteristics that can be easily assessed based on charter applications, including school themes, exclusionary discipline policies, policies on midyear enrollment, and plans to engage with the community/parents. By providing answers to these underlying factors, researchers can better explain the vast variation in charter school effectiveness, helping policymakers better differentiate and anticipate the effectiveness of schools in their decision-making process of authorizing new charter schools.

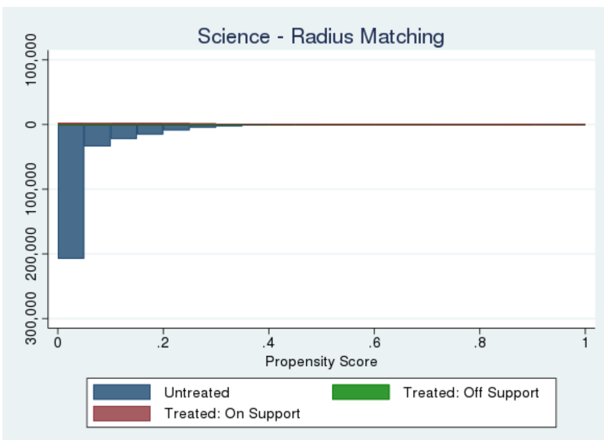
Appendix

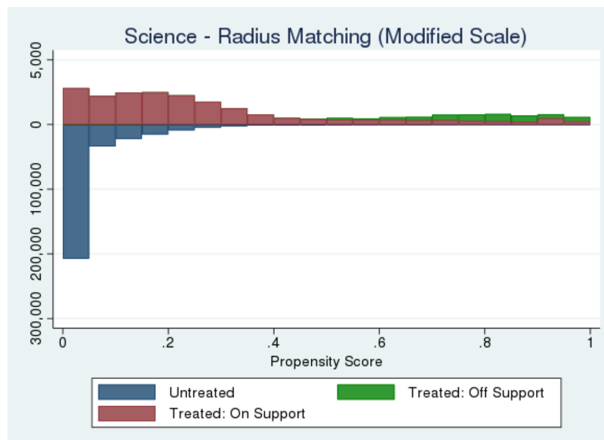
Figure A1. Propensity Score Matching – Additional Common Support Graphs

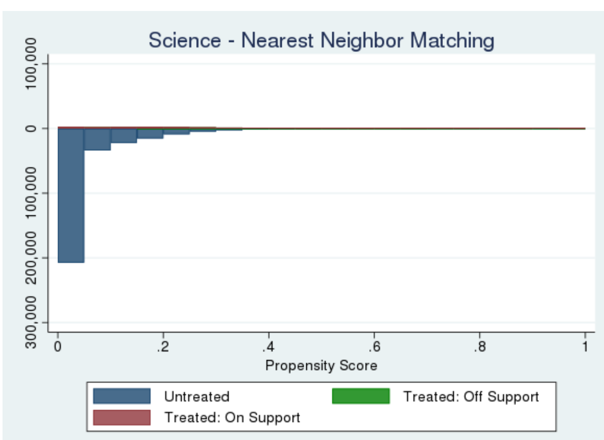


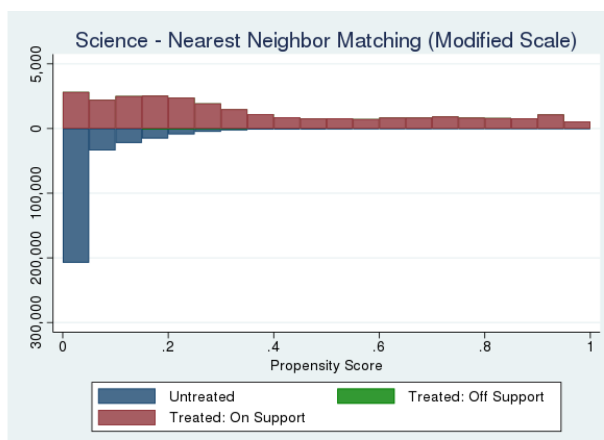

Math - Nearest Neighbor Matching
100,000
0
100,000
200,000
300,000
0
.2
.4
.6
.8
1
Propensity Score
Untreated
Treated: Off Support
Treated: On Support


Math - Nearest Neighbor Matching (Modified Scale)
5,000
0
100,000
200,000
300,000
0
.2
.4
.6
.8
1
Propensity Score
Untreated
Treated: Off Support
Treated: On Support


Science - Radius Matching
100,000
0
100,000
200,000
300,000
0
.2
.4
.6
.8
1
Propensity Score
Untreated
Treated: Off Support
Treated: On Support


Science - Radius Matching (Modified Scale)
5,000
0
100,000
200,000
300,000
0
.2
.4
.6
.8
1
Propensity Score
Untreated
Treated: Off Support
Treated: On Support


Science - Nearest Neighbor Matching
100,000
0
100,000
200,000
300,000
0
.2
.4
.6
.8
1
Propensity Score
Untreated
Treated: Off Support
Treated: On Support


Science - Nearest Neighbor Matching (Modified Scale)
5,000
0
100,000
200,000
300,000
0
.2
.4
.6
.8
1
Propensity Score
Untreated
Treated: Off Support
Treated: On Support

References

Abdulkadiroğlu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. The Quarterly Journal of Economics, 126(2), 699-748.

Angrist, J. D., Dynarski, S. M., Kane, T. J., Pathak, P. A., & Walters, C. R. (2010). Inputs and impacts in charter schools: KIPP Lynn. American Economic Review, 100(2), 239-243.

Angrist, J. D., Pathak, P. A., & Walters, C. R. (2013). Explaining charter school effectiveness. *American Economic Journal: Applied Economics*, *5*(4), 1-27.

Ballou, D., Teasley, B., & Zeidner, T. (2006). Charter schools in Idaho. In Prepared for the National Conference on Charter Schools Research, Vanderbilt University, September (Vol. 29).

Batdorff, M., Maloney, L., May, J. F., Speakman, S. T., Wolf, P. J. & Cheng, A. (2014). Charter School Funding: Inequity Expands. University of Arkansas – Department of Education Reform: School Choice Demonstration Project. Retrieved from http://www.uaedreform.org/wp-content/uploads/charter-funding-inequity-expands.pdf

Bean, M. (2010). Dewey to Delpit: The No-Excuses Charter School Movement. Retrieved from http://edcommentary.blogspot.com/p/no-excuses-charter-movement.html

Bellafante, G. (2017). A Brooklyn Charter School Looks Past 'No Excuses'. *The New York Times*. Retrieved from https://www.nytimes.com/2017/03/09/nyregion/brooklyn-charter-school-discipline.html

Berends, M., Watral, C., Teasley, B., & Nicotera, A. (2008). Charter school effects on achievement: Where we are and where we're going. Charter school outcomes, 243-266.

Bettinger, E. P. (2005). The effect of charter schools on charter students and public schools. Economics of Education Review, 24(2), 133-147.

Betts, J. R. (2006). *Does School Choice Work?: Effects on Student Integration and Achievement*. Public Policy Institute of California.

Bifulco, R. (2012). Can Nonexperimental Estimates Replicate Estimates Based on Random Assignment in Evaluations of School Choice? A Within-Study Comparison. *Journal of Policy Analysis and Management*, *31*(3), 729-751.

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

Booker, K., Gilpatric, S. M., Gronberg, T., & Jansen, D. (2007). The impact of charter school attendance on student performance. Journal of Public Economics, 91(5), 849-876.

Carr, S. (2014) How Strict is Too Strict? *The Atlantic*. Retrieved from https://www.theatlantic.com/magazine/archive/2014/12/how-strict-is-too-strict/382228/

Carruthers, C. K. (2012). New schools, new students, new teachers: Evaluating the effectiveness of charter schools. Economics of Education Review, 31(2), 280-292.

Carter, S. C. (2000). *No Excuses: Lessons from 21 High-Performing, High-Poverty Schools*. Heritage Foundation, 214 Massachusetts Avenue, NE, Washington, DC 20002.

Center for Research on Education Outcomes [CREDO]. (2009). Multiple choice: Charter school performance in 16 states.

Cheng, A., Hitt, C., Kisida, B., & Mills, J. N. (2017). "No Excuses" Charter Schools: A Meta-Analysis of the Experimental Evidence on Student Achievement. Journal of School Choice, 1-30.

Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of policy analysis and management*, *27*(4), 724-750.

Cremata, E., Davis, D., Dickey, K., Lawyer, K., Negassi, Y., Raymond, M., & Woodworth, J. L. (2013). National charter school study. *Center for Research on Education Outcomes, Stanford University*.

Csuvarszki, E. (2016). How Have "No Excuses Schools" Appeared and Changed Over Time in the US? Retrieved from http://commons.trincoll.edu/edreform/2016/05/how-have-no-excuses-schools-appeared-and-changed-over-time-in-the-us/

Curto, V. E., & Fryer Jr, R. G. (2014). The potential of urban boarding schools for the poor: Evidence from SEED. *Journal of Labor Economics*, *32*(1), 65-93.

Disare, M. (2016). 'No excuses' no more? Charter schoosl rethink discipline after focus on tough consequences. *Chalkbeat*. Retrieved from http://www.chalkbeat.org/posts/ny/2016/03/07/no-excuses-no-more-charter-schools-rethink-discipline-after-focus-on-tough-consequences/

Dobbie, W., Fryer, R. G., & Fryer Jr, G. (2011). Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone. American Economic Journal: Applied Economics, 3(3), 158-187.

Dobbie, W., & Fryer Jr, R. G. (2013). Getting beneath the veil of effective schools: Evidence from New York City. American Economic Journal: Applied Economics, 5(4), 28-60.

Dobbie, W. S., & Fryer Jr, R. G. (2016). Charter Schools and Labor Market Outcomes (No. w22502). National Bureau of Economic Research.

Eberts, R. W., & Hollenbeck, K. (2001). An examination of student achievement in Michigan charter schools.

Ellison, S. (2012). It's in the name: A synthetic inquiry of the Knowledge is Power Program [KIPP]. *Educational Studies*, *48*(6), 550-575.

Garland, S. (2016). The end of "no excuses" education reform? *The Hechinger Report*. Retrieved from http://hechingerreport.org/the-end-of-no-excuses-education-reform/

Glazer, J. L., & Egan, C. (2016). The Tennessee Achievement School District: Race, history, and the dilemma of public engagement. *Tennessee Consortium of Research, Evaluation, and Development. Retrieved March*, *6*, 2016.

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

Gleason, P., Clark, M., Tuttle, C. C., & Dwoyer, E. (2010). The Evaluation of Charter School Impacts: Final Report. NCEE 2010-4029. National Center for Education Evaluation and Regional Assistance.

Golann, J. W. (2015). The paradox of success at a no-excuses school. Sociology of education, 88(2), 103-119.

Goodman, J. F. (2013). Charter management organizations and the regulated environment: Is it worth the price?. Educational Researcher, 42(2), 89-96.

Hansen, B. B. (2004). Full matching in an observational study of coaching for the SAT. *Journal of the American Statistical Association*, *99*(467), 609-618.

Hanushek, E. A., Kain, J. F., Rivkin, S. G., & Branch, G. F. (2007). Charter school quality and parental decision making with school choice. Journal of public economics, 91(5), 823-848.

Hastings, J. S., Neilson, C. A., & Zimmerman, S. D. (2012). The effect of school choice on intrinsic motivation and academic outcomes (No. w18324). National Bureau of Economic Research.

Henry, G. T., & Yi, P. (2009). Design matters: A within-study assessment of propensity score matching designs. *Association for Public Policy and Management, Washington DC*.

Henry G., Zimmer, R., Attridge, J., Kho, A. & Viano, S. (2015). Teacher and student migration in and out of Tennessee's Achievement School District. Retrieved from *https://peabody.vanderbilt.edu/research/tnedresearchalliance/files/ASD_Teacher_Student_Migration.pdf*

Henry, G., Zimmer, R., Kho, A. & Pham, L. (2017) Recruitment and Retention of Teachers in Tennessee's Achievement School District and iZone Schools. Retrieved from https://peabody.vanderbilt.edu/research/tnedresearchalliance/files/Teacher_Retention_in_ASD_and_iZone_vF.pdf

Horn, J., & Wilburn, D. (2013). *The mismeasure of education.* Charlotte, NC: Information Age Publishing.

Hoxby, C. M., Murarka, S., & Kang, J. (2009). How New York City's charter schools affect achievement. Cambridge, MA: New York City Charter Schools Evaluation Project, 1-85.

Hoxby, C. M., & Rockoff, J. E. (2004). The impact of charter schools on student achievement. Cambridge, MA: Department of Economics, Harvard University.

Imberman, S. A. (2011). The effect of charter schools on achievement and behavior of public school students. Journal of Public Economics, 95(7), 850-863.

Kretzmann, J., & McKnight, J. P. (1996). Assets-based community development. *National Civic Review*, *85*(4), 23-29.

Lack, B. (2009). No Excuses: A Critique of the Knowledge Is Power Program (KIPP) within Charter Schools in the USA. *Journal for Critical Education Policy Studies (JCEPS)*, *7*(2).

Lake, R., Bowen, M., Demeritt, A., McCullough, M., Haimson, J., & Gill, B. (2012). Learning from Charter School Management Organizations: Strategies for Student Behavior and Teacher Coaching. *Mathematica Policy Research, Inc.*

Lauen, D. L., Fuller, B., & Dauter, L. (2014). Positioning charter schools in Los Angeles: Diversity of form and homogeneity of effects. American Journal of Education, 121(2), 213-239.

Lavertu, S., & Witte, J. (2008). The impact of Milwaukee charter schools on student achievement. *Policy*.

Miron, G., Cullen, A., Applegate, B., & Farrell, P. (2007). Evaluation of the Delaware charter school reform: Final report.

National Alliance for Public Charter Schools [NAPCS]. (2016). A Closer Look at the Charter School Movement: Charter Schools, Students, and Management Organizations, 2015-16. Retrieved from http://www.publiccharters.org/wp-content/uploads/2016/02/New-Closed-2016.pdf

Nicotera, A., Mendiburo, M., & Berends, M. (2009). Charter school effects in an urban school district: an analysis of student achievement gains in Indianapolis. *Prepared for School*

Choice and School Improvement: Research in State, District and Community Contexts, Vanderbilt University.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*(3), 169-188.

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

Shadish, W. R., Clark, M. H., & Steiner, P. M. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association*, *103*(484), 1334-1344.

Solmon, L. C., & Goldschmidt, P. (2004). Comparison of traditional public schools and charter schools on retention, school switching, and achievement growth.

Stuart, E. A., & Green, K. M. (2008). Using full matching to estimate causal effects in nonexperimental studies: examining the relationship between adolescent marijuana use and adult outcomes. *Developmental psychology*, *44*(2), 395.

Thernstrom, A., & Thernstrom, S. (2004). *No excuses: Closing the racial gap in learning*. Simon and Schuster.

Tuttle, C. C., Gill, B., Gleason, P., Knechtel, V., Nichols-Barrer, I., & Resch, A. (2013). KIPP Middle Schools: Impacts on Achievement and Other Outcomes. Final Report. Mathematica Policy Research, Inc.

U. S. Department of Education, National Center for Education Statistics [NCES]. (2017). Table 216.20: Number and enrollment of public elementary and secondary schools, by school level, type, and charter and magnet status: Selected years, 1990-91 through 2015-16. In U.S. Department of Education, National Center for Education Statistics (Ed.), *Digest of Education Statistics* (2016 ed.). Retrieved from https://nces.ed.gov/programs/digest/d17/tables/dt17_216.20.asp

Whitman, D. (2008). *Sweating the Small Stuff: Inner-City Schools and the New Paternalism*. Thomas B. Fordham Foundation & Institute. 1701 K Street NW Suite 1000, Washington, DC 20006.

Witte, J., Weimer, D., Shober, A., & Schlomer, P. (2007). The performance of charter schools in Wisconsin. Journal of Policy Analysis and Management, 26(3), 557-573.

Zimmer, R., & Buddin, R. (2006). Charter school performance in two large urban districts. *Journal of Urban Economics*, *60*(2), 307-326.

Zimmer, R., Gill, B., Booker, K., Lavertu, S., Sass, T. R., & Witte, J. (2009). Charter schools in eight states: Effects on achievement, attainment, integration, and competition (Vol. 869). Rand Corporation.

Zimmer, R., Henry, G. & Kho, A. (2017) The Effects of School Turnaround in Tennessee's Achievement School District and Innovation Zones. *Educational Evaluation and Policy Analysis*, 0162373717705729.