Three Essays on Economic Access to Public K-12 Education

Leah R. Clark

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Economics and Public Policy

Carnegie Mellon University

March 26, 2019

Accepted by the Dissertation Committee and Approved by the Deans:

Dennis Epple (Co-chair): ______________________ Date: ________

Lowell Taylor (Co-chair): ______________________ Date: ________

Brian Junker: ______________________ Date: ________

Lindsay Page: ______________________ Date: ________

Ramayya Krishnan: ______________________ Date: ________

Dean, Heinz College

Robert Dammon: ______________________ Date: ________

Dean, Tepper School of Business

Three Essays on Economic Access to Public K-12 Education

Leah R. Clark

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Economics and Public Policy

Carnegie Mellon University

March 26, 2019

Dissertation Committee:

Dennis Epple

Lowell Taylor

Brian Junker

Lindsay Page

Disclaimer

The research reported here was supported, in whole or in part, by the Institute of Education Sciences, U.S. Department of Education, through grant R305B150008 to Carnegie Mellon University. The opinions expressed are those of the author and do not represent the views of the Institute or the U.S. Department of Education.

Dedication

For Joanne & Marilyn

Acknowledgements

My co-advisors, Dennis Epple and Lowell Taylor, for job market and research support, as well as my committee members, Lindsay Page and Brian Junker. My dad, Darryl, for everything & always. My brothers, Ben and Nat, and sisters-in-law, Mary and Carolyn, for the stream of adorable and distracting babies. My non-PhD friends (esp. Keisha Gayle-Luz, Kimberly-Anne Moriarty, Hilary Lawlor, Maria Figliola, and Ben Howard) for much-needed perspective and fun. My Heinz & Tepper cohorts for endless support (esp. Erica VanSant, Nam Ho, and Wenting Yu during Qualidays). Lynn Conell-Price and Anna Mayo for job market solidarity. My original mentor, Tandayi Jones, and my PhD gurus Matt Eisenberg, Melanie Zaber, and Thomas Goldring for endless wisdom. CMU's kind and helpful administrators (Gretchen Hunter, Michelle Wirtz, Lawrence Rapp, Laila Lee, and Audrey Russo). And Cornelius Mduduzi Masuku for meeting me at the finish line.

Abstract

85% of children in the U.S. attend public schools for their primary and secondary education. Variation in education and labor market outcomes across schools is substantial, and access to high-quality public schools depends largely on where children live. Yet there is more to educational access than which school district a child calls home. Parents and caregivers play a further role through local school enrollment choices and in the supplementary resources and effort they provide. Government actors – from teachers and caseworkers to superintendents and state-level policymakers – implement policies that augment both the way children access schools and the educational resources schools offer. This dissertation is comprised of three essays that consider how children access public education, and what factors mediate their experiences.

Adverse home and family circumstances impact children's access to education, but typical administrative data cannot identify these disadvantages beyond a simple household income proxy (e.g., free or reduced-price lunch eligibility). In the first essay, "Disadvantage Beyond Poverty: Adverse Childhood Experiences, School Choice, and Educational Outcomes," I leverage human services data to identify more severely disadvantaged children: those linked to child welfare investigations prior to kindergarten (one-in-eight children in the city under study). Regardless of which elementary school they attend, these children miss 25% more school days than non-disadvantaged children and have suspension odds that are 57% higher. These gaps are significantly larger than those for low-income children not linked to investigations. Moreover, children linked to child welfare investigations are systematically less likely to enroll in charter or magnet schools over traditional public schools, in contrast to other low-income children who are nearly as likely as their counterparts to enroll in charter or magnet schools. Thus, in aggregate at the school level, a handful of traditional public schools disproportionately enroll the most disadvantaged students. By carefully controlling for these sorting patterns and leveraging measures of early childhood disadvantage determined prior to school entry, I recover causal peer effect estimates, which show that having more disadvantaged peers significantly increases students' own suspension probabilities.

The second essay, "An Intervention to Reduce Chronic Absenteeism," offers a quasi-experimental evaluation of a truancy prevention program that was piloted in two urban K-8 schools during the 2013-14 school year. Using a triple-differences specification, I find that the program was successful in reducing chronic absenteeism rates among persistently low-income students, but the treatment

effect is only marginally significant. These results suggest that encouraging school attendance and providing resources to address attendance barriers can improve the educational access of low-income students via a reduction in chronic absenteeism.

In the third essay, "Persistent and Wide-Ranging Differences in the Income and Racial Segregation of Children," I document the income and racial segregation experienced by children in neighborhoods and schools throughout the nation. Comparing segregation estimates across commuting zones, I show that racial and income segregation exhibit distinct geographical patterns. Racial segregation operates across school district boundaries, while income segregation persists within school districts. Of the demographic variation experienced by children in schools, more than 40% of variation in low-income demographics occurs within school districts, but only 18% of the variation in non-white demographics occurs within school districts. These nuances are relevant to policies that may impact how children of different racial and economic backgrounds access public education, whether directly or through spillover effects.

Contents

1 Disadvantage Beyond Poverty: Adverse Childhood Experiences, School Choice, and Educational Outcomes **3**

1.1 Introduction 3

1.2 Data 7

1.2.1 Human service variables 9

1.2.2 School variables 11

1.3 Selection into schools 13

1.3.1 Methodology 14

1.3.2 Results 15

1.4 Outcomes Model 17

1.4.1 Identification & Methodology 20

1.4.2 Results 23

1.4.3 Results by race 29

1.5 Discussion 31

2 An Intervention to Reduce Chronic Absenteeism **42**

2.1 Introduction 42

2.2 Background & Data 44

2.2.1 Data 46

2.3 Methodology 48

2.4 Results 51

2.5 Discussion 54

3 Persistent and Wide-Ranging Differences in the Income and Racial Segregation of Children **55**

3.1 Introduction . . . 55
3.2 Data & Geographies . . . 58
3.3 Measures of Segregation . . . 61
3.3.1 Dissimilarity Index . . . 62
3.3.2 Rank-order Variance Ratio Segregation Index . . . 63
3.3.3 Segregation estimates . . . 64
3.4 Methodology . . . 66
3.5 Findings . . . 69
3.6 Discussion . . . 73

Chapter 1

Disadvantage Beyond Poverty: Adverse Childhood Experiences, School Choice, and Educational Outcomes

1.1 Introduction

Economic disadvantage has far-reaching implications for parental investments in children's human capital development and children's own educational outcomes (Cunha & Heckman 2007; Dahl & Lochner 2012).[1] Of course, not all low-income children and families are the same, and poverty may not be the direct cause of low parental investments or adverse outcomes. For some, poverty comes alongside physical, mental, and environmental health issues that worsen educational outcomes, even (or perhaps especially) when exposure comes very early in life (Case et al. 2004; Aizer et al. 2009; Aizer et al. 2018). Similarly, child maltreatment – another correlate of poverty and parental investments – is connected to worse educational outcomes via survey data (Pieterse 2014; Slade & Wissow 2007). While rich administrative school data are available to researchers, they typically contain few measures of children's family and early childhood experiences, limiting the ability of researchers to understand how the nuances of disadvantage factor into educational outcomes.

[1] I will use "parent" or "parents" to refer to the adult caretakers that make decisions regarding children's educational experiences. I will also use the singular pronoun "their/them."

This paper leverages a unique dataset that links information on early childhood human services exposure to data on school enrollments and behavioral outcomes. These data enable analyses of the educational experiences of children linked to child welfare investigations early in life (12% of kindergartners in the city under study) relative to those who were low-income early in life but never linked to child welfare (52% of kindergartners). These analyses produce several new findings on the relationship between adverse early childhood experiences and educational outcomes. (For brevity, I will refer to students linked to child welfare investigations prior to kindergarten as Early Child Welfare students). First, Early Child Welfare students are less than one-half as likely to enroll in choice schools (in lieu of neighborhood public schools) relative to children without child welfare ties. This stands in contrast to low-income children, who are only slightly less likely to enroll in magnet and charter schools than their non-low-income counterparts. Second, while low-income students have significantly higher rates of absence and suspension than non-low-income students, these disparities are even larger for Early Child Welfare students. Regardless of which school they attend, Early Child Welfare students miss 25% more school days and are 57% more likely to be suspended than other students.

Moreover, I find evidence of significant adverse peer effects on suspension. Students who have more disadvantaged peers – whether Early Child Welfare, low-income, or both – are themselves more likely to be suspended. Since Early Child Welfare students are disproportionately concentrated in a subset of traditional public schools (and not in magnet or charter schools), so, too, are these adverse spillover effects.

Why should early child welfare investigations provide insight into later educational outcomes? The most literal interpretation of being an Early Child Welfare student is that, at some point before kindergarten, someone made a phone call to a hotline to report concerns of abuse or neglect concerning the student, or another child who lives with them. Parental risk factors for child maltreatment include drug and alcohol abuse, intimate partner violence, and criminal justice involvement (Doyle & Aizer 2018). Potentially, the parents of Early Child Welfare students make lower investments, on average, in their child's wellbeing in early childhood, and this correlates with later disparities in educational investments. This interpretation is consistent with the findings of this paper: parents of children linked to child welfare may be less likely to invest the time and effort to get their children enrolled in choice schools and less likely to facilitate regular school attendance for their children. These disparities do not necessarily reflect mal-intent: they could also be attributable to differences in parental resources (financial resources but also time and health) or preferences and savvy with respect to education.

From the child's perspective, child welfare investigations may indicate exposure to abuse, neglect, or living in a resource-poor family, among other factors. Research on later-in-life outcomes for maltreated children find decreased earnings, more criminal activity, and a continuing cycle of violence (Currie & Spatz Widom 2010; Currie & Tekin 2012). Perhaps, then, it is unsurprising that Early Child Welfare students face significantly higher odds of suspension, regardless of which school they attend. Hopefully, child welfare investigations are effective in improving children's circumstances, but the experience of the investigation itself may also have an adverse impact. Approximately one-in-five investigations result in the child being removed from the home (potentially to a relative's house) for at least one day.

Critically for the causal framework employed in this paper, Early Child Welfare can be measured prior to starting school. Children are linked to child welfare investigations as early as birth. This is true of other human services variables, as well, enabling the use of early childhood Medicaid enrollments as a proxy for living in a low-income household prior to school entry. Most children who use Medicaid enroll in the year of their birth, and take-up among young children is high. Since enrollment is means-tested, this provides a reliable proxy for whether children live in low-income families in early childhood (Kenney et al. 2012). (For brevity, I will refer to students who were enrolled in Medicaid prior to kindergarten as Early Medicaid students). In contrast to Early Child Welfare and Early Medicaid, variables commonly available in school data – like indicators for whether a student receives free/reduced-price lunch or special education services – are not necessarily indicative of early childhood factors and may be influenced by policies and practices that vary across schools and over time (Domina et al. 2018; Michelmore & Dynarski 2017).

Whether disadvantaged children have equal access to choice schools is a major concern for education reform efforts that center around expanding school choice options in order to improve access to quality education for children living in low-performing school districts (Epple & Romano 1998). To investigate whether Early Child Welfare and Early Medicaid students are equally as likely to enroll in magnet, charter, or private schools (in lieu of traditional public schools) as their counterparts, I implement a multinomial choice model of the type of school students first enroll in for kindergarten as a function of Early Child Welfare and Early Medicaid (in addition to gender and race controls). To account for potential neighborhood effects that operate prior to school enrollment, the analysis includes zip code fixed effects, which offer reasonable approximations of neighborhoods in the city under study. Interestingly, zip code fixed effects have little-to-no impact on the estimates of interest, suggesting either Early Child Welfare and Early Medicaid operate orthogonally to zip code effects, or they explain a substantial portion of zip code effects, at least as they pertain to educational enrollments.

The findings of the selection model – that disadvantaged children are systematically less likely to enroll in alternatives to traditional public school – are new to the school selection literature, but do not refute current wisdom. Conditional on income, private school voucher decliners are more likely to be single mothers working full-time, suggesting that non-income family resources (like parental time) play an important role in accessing choice options (Cowen, 2010). Studies on the role of race in school choice selection suggest that choice schools may increase racial segregation by disproportionately enrolling students of a particular race relative to local district demographics (Bifulco & Ladd, 2006; Garcia, 2008). My model indicates that it is white students in this district who are disproportionately enrolling in charter schools. While the analysis does not consider special needs students directly, to the extent that Early Child Welfare students are more likely to receive special education services in school, their enrollment disparity is also in line with ongoing concerns that choice schools do not serve special needs students proportionally (Cowen & Winters, 2013).

These findings raise questions about whether the disparities in educational access of Early Child Welfare students extend to school outcomes. And, if their outcomes are worse, does their presence in the classroom cause other students to have worse outcomes, as well? In the second major analysis of this paper, I use data on elementary school attendance and suspension outcomes to address these questions. This paper is not the first to use administrative measures of child welfare investigations as a risk factor pertinent to educational outcomes, but those analyses use measures that are potentially endogenous to schools and are, thus, limited in their causal claims (Fantuzzo et al. 2011; Fantuzzo et al. 2014; Ryan et al. 2018).

Absences and suspensions are of particular interest in this context for several reasons. First, they are understudied in the economics literature, but play a critical role in student learning (Gershenson et al. 2017). Second, comprehensive human services histories are primarily available for younger students, many of whom are not yet taking comprehensive standardized exams (which start in third grade). Finally, these outcomes offer a useful contrast. For elementary school students, attendance is primarily a parent's choice, while suspension reflects a student's behavior in combination with how the school decided to respond to that behavior. These variables are not simply outcomes, but also measures of inputs to the education production function (Lazear, 2001). Thus, disparities in attendance provide insight into the challenges that schools face in effectively educating students.

Identifying the impact of Early Child Welfare and Early Medicaid on outcomes requires separating multiple pathways through which the effects could operate (directly and through peers) and multiple ways in which these pathways could be confounded (school policies and sorting). Fortunately, having pre-determined measures of children and their peers ensure that the independent variables of interest cannot be influenced by educational outcomes. This enables an approach like that of Carrell & Hoekstra (2010), who study the impact of domestic violence exposure (as proxied by protective orders) on educational outcomes. As with Early Child Welfare and Early Medicaid, students have no influence over whether their peers are exposed to domestic violence, so it offers a source of exogenous variation in peer group composition across schools and grades, after accounting for the challenge to identification stemming from nonrandom selection into schools.

Since this paper provides evidence of selection into schools along dimensions of early disadvantage, it is reasonable to assume there are other factors influencing school selection that are not captured in the data. Such unobserved variables create a challenge for identification: if students' outcomes at a certain school are correlated, does that represent the school's impact, or does that reflect the fact that students (or parents) at the school are similar? Thus, identifying the causal effect of early childhood human services exposure on absences and suspensions requires carefully controlling for sorting. This is done by employing a number of fixed effects. Ultimately, the identifying variation in the outcomes regressions occurs within school-grade cohorts over time, and separately from district- or school-level trends.

Analyses will repeatedly return to the question of whether a simple control for poverty can account for heterogeneity in the experiences of disadvantaged children. Throughout the analyses, I find evidence that the Early Child Welfare variable identifies children who experience worse outcomes, on average, than those who are "just" Early Medicaid. Moreover, Early Child Welfare students exert a unique impact on their peers.

1.2 Data

The data stem from a partnership between a midsize urban school district and the human services department that serves the city and the surrounding county. Child welfare and Medicaid histories, school enrollments, and attendance and suspension outcomes are observed for every student who enrolls in kindergarten in the district's traditional public schools or magnet schools from the 2012-

	Selection model	Outcomes model
Students	8,502	20,020
School years	3	4
Grades	K	K-5
School types	All major	Trad. & magnet
Avg. obs. per student	1	2.2 (1.1)
Female	48.1%	48.7%
White	38.9%	31.4%
Black	45.1%	52.7%
Other minority	16.0%	15.9%
Early child welfare	12.4%	16.6%
Early medicaid	63.3%	66.2%
ECW & EM	11.1%	14.5%

Standard deviation in parentheses.

Figure 1.1: Summary statistics.

13 school year through the 2016-17 school year.[2] From the 2013-14 school year forward, human service histories and enrollments are additionally observed for children who live in the city but enroll outside of the district – primarily in charter or private schools. However, attendance and suspension outcomes are never observed for non-district schools.

The selection and outcomes models (described in subsequent sections) rely on different but overlapping cuts of the dataset. Figure 1.1 offers a comparison of the two subsamples. The reasons for the data restrictions are explained in the methodology descriptions of each model, but, in short, the selection model utilizes kindergarten enrollments from 2013-14 on (when the universe of schooling options is observed), while the outcomes model makes use of data for all elementary grades from 2012-13 forward, but is limited to traditional public and magnet schools. Thus, to understand what demographics are typical of the city's children in recent years, one should focus on the selection model statistics.

Approximately 2,800 children in the city enroll in kindergarten for the first time each year (see Figure 1.2). 56% enroll in traditional public schools, 18% in the district's magnet schools, 9% in non-district-affiliated public charter schools, and 15% in private schools. A small number of children

[2]The 2015-16 school year is excluded from the main analyses due to a censoring issue in the enrollment data for that year.

School type	2013-14	2014-15	2016-17
Traditional public	1,794	1,603	1,373
Magnet	527	540	486
Charter	241	268	272
Private	460	481	340
Other*	40	36	41
Total	3,062	2,928	2,512

*Includes day cares, special needs schools, and home school.

Figure 1.2: Kindergarten enrollment by school type.

enroll in kindergarten in day cares or special needs schools (e.g., schools for the blind). Very few children are enrolled in home school for kindergarten, but it is possible that some home-schooled children are left out of the data system. On average, 45% of enrolling kindergartners are identified as black, 39% white, and 9% multi-racial, with the remaining 7% identified as Asian, Hispanic, or Native American. There are not many English Language Learners in the district.

1.2.1 Human service variables

Human service histories in the dataset are comprised of a set of indicators summarizing whether a child interacted with a given service area during a given school year – even if the child was not yet enrolled in school. Consider a child enrolling in kindergarten for the first time in Fall 2013 after their 5th birthday: the data indicate which services they were associated with in 2012-13 (approximately age 4), 2011-12 (approximately age 3), and so on. I will use "early childhood" to denote the years between birth and beginning kindergarten (typically between ages 5 and 6). Early childhood human service histories are missing only for children who did not live inside the county prior to enrolling in a city school. Unfortunately, the data do not differentiate between students who never used human services and students who lived outside of the county.

This paper focuses on two aspects of student's early childhood human service histories: child welfare investigations and Medicaid enrollments. "Early Child Welfare" or "ECW" and "Early Medicaid" or "EM" will be used to denote indicator variables for any involvement in the respective service in early childhood. Thus, an "ECW student" is a student who was linked to at least one child welfare investigation prior to starting kindergarten, and an "EM student" is a student who was enrolled in Medicaid at least once prior to starting kindergarten. Medicaid is a means-tested program, with high take-up among children. Like free/reduced lunch eligibility, it is often used as a proxy for children living in low-income households. As with any voluntary program, concerns about selection into the program are reasonable. The population of Early Medicaid students, however, is fairly consistent as children age, with the vast majority enrolled in Medicaid every year between birth and kindergarten.

Most ECW students, on the other hand, are only associated with a child welfare investigation during one or two years (the investigations themselves are typically much shorter than a year). Involvement begins at any age, though more than half of ECW students are linked to their first investigation before their second birthday. Child welfare investigations typically arise after someone contacts a hotline to report suspicions of abuse or neglect concerning a particular child. Children who live with children referred for investigation are also referred. Many professionals who work with children are mandated reporters – that is, the law requires them to report suspicions of abuse or neglect within a strict time window. Mandated reporters include doctors and child care workers.

Child welfare investigations sometimes result in a child being removed from the home and placed in foster care, often with a relative. Approximately one-in-five ECW students are placed in foster care for some period of time in early childhood. Other ECW students receive supports or services to ensure their well-being while remaining in their homes.

Participation in child welfare is involuntarily, so there is no concern that ECW reflects selection on the part of parents or children. However, it is possible that some children who should have been referred to child welfare in early childhood were not. It is also possible that the referring party is biased in their choice to make a report. The latter concern has led to a number of studies investigating why black children are referred to child welfare at much higher rates than white children. These studies are not able to rule out racial bias, but they do present some potential alternative explanations. For example, Maloney et al. (2017) use birth records linked to administrative data on child welfare referrals and find that differences in marital status and maternal age fully explain the racial gap in referrals. Again, this cannot rule out bias, but it makes sense that resource limitations or limited parental experience might give rise to circumstances that prompt reports to child welfare. Conditional on referral to child welfare, black and white children are equally likely to be placed in foster care.

In sum, ECW students may have encountered a host of negative experiences before starting kindergarten, but there is no obvious way to model what this indicator should mean in terms of human capital formation. These children may also have been impacted – positively or negatively – by well-intentioned services and providers, including foster care placements. The simplest interpretation of the indicator is that, at some point early in a child's life, someone expressed concern about their well-being. Despite the variable's murky implications, ECW provides a useful signal about early childhood experiences.

1.2.2 School variables

School data is compiled at the student-enrollment-year level; that is, for each student and each school in which they enroll, I observe the school, grade, enrollment/withdrawal dates, and number of days absent or suspended. Additionally, I observe some demographic information, like race, gender, age at start of each year, and zip code of residence. Indicators of special need and free/reduced lunch eligibility are also in the data, but will not be used in subsequent analyses, since they may change over time and in response to school practices.

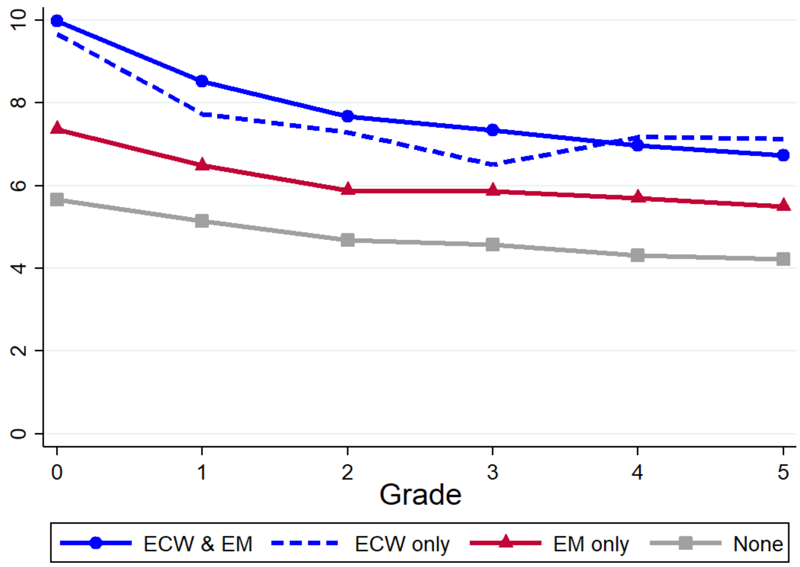
The dependent variables in the outcomes model are absence rates and suspensions. Absence rates are calculated as days missed – net of suspension days – as a share of total days enrolled.[3] Most children are not suspended, and those that are typically serve just one day of suspension, so the suspension outcome is simply an indicator for whether a child was suspended during a particular enrollment-year. Enrollments are excluded if they (a) last 50% of the school year or less, or (b) exhibit an absence rate of 75% or more. The latter exclusion reflects the concern that these records reflect administrative errors rather than real school enrollments.

Figure 1.3 shows average absence and suspension outcomes by grade for students with various human service histories. These data come from the outcomes model dataset, so 14% of students have both Early Child Welfare and Early Medicaid exposure; 2% of students have Early Child Welfare exposure without Early Medicaid exposure; 52% of students have Early Medicaid exposure without Early Child Welfare exposure; and 32% of students have neither. This latter group has the best average outcomes, while the "EM only" group has somewhat worse outcomes, on average, and the two ECW groups have the worst average outcomes of all.

The most striking observation from these figures is how closely the "ECW only" lines track the "ECW & EM" lines. This suggests that the Early Child Welfare measure – not the Early Medicaid measure – is more strongly associated with adverse outcomes. Of course, the outcomes model will explore this contention more rigorously.

[3] I do not differentiate between unexcused and excused absences out of a concern that the distinction is endogenous. For the same number of missed days, low-income children have fewer excused absences than do non-low-income children.

A. Absence rate (% of days)



B. Suspension rate

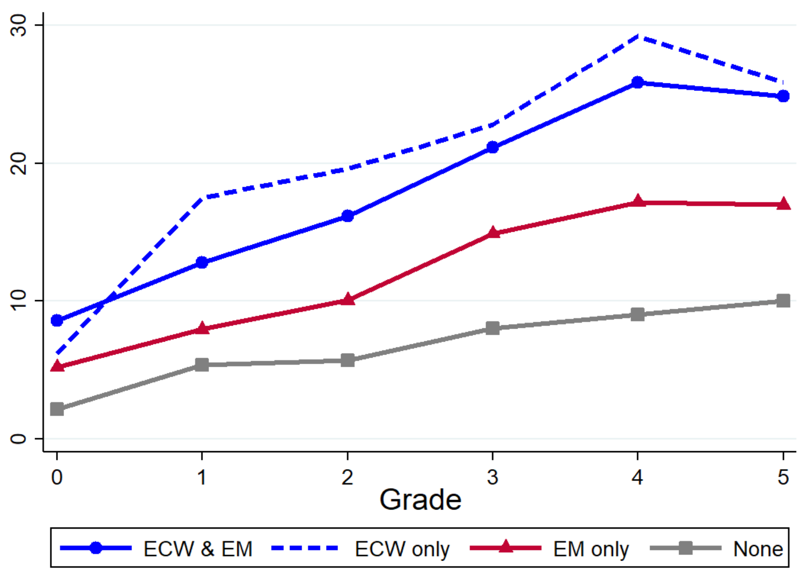


Figure 1.3: District average outcomes by grade and Early Child Welfare/Early Medicaid exposure.

Overall, absence rates decline slightly as students ascend grades. Parents may take school attendance more seriously as students get older. Suspension rates increase substantially as students ascend grades, though approximately 5% of kindergartners receive suspensions – enough to include kindergartners in the suspension outcomes model. The outcomes model will include district-level grade-by-year fixed effects to separate out these general trends from variation attributable to particular schools, cohorts, or students.

1.3 Selection into schools

School selection can occur any time during a child's education, but most children enroll in kindergarten in a particular elementary school and stay in that school through fifth grade. Since the human services data permit characterization of children prior to school entry, I will focus on modeling school selection at the point of first entry into elementary school. It is possible that parents will change their mind about their child's school and switch them into a different school type later on. However, later school selection could occur based on specific school experiences unobserved to the econometrician (e.g., a child being bullied), while initial school selection more likely reflects general expectations for a school, and those expectations may be shared across many parents.

What does a child's first kindergarten enrollment say about their parents' preferences? Choice schools (including private schools) likely require more advance planning – oversubscribed choice options typically require applications[4] more than six months prior to the start of school – but some will admit students through the start of the school year. Enrollment in a neighborhood public school could reflect a passive choice (parents did not consider alternatives), an active choice (parents searched for alternatives and concluded it was the best choice), or a temporary choice (parents are still searching for alternatives, but want their child to start school on time). Thus enrollment does not always reflect parents' preferences for schools, but for parents making the passive choice, enrollment in traditional public schools is practically guaranteed.

So if some degree of heterogeneity among parents is captured by which schools they select for their children in kindergarten, how does this heterogeneity extend to children? Is there any evidence that children enrolled in traditional public schools differ from those enrolled in choice schools? This section puts forth a multinomial logistic choice model of school type for first kindergarten enrollment

[4]To my knowledge, no schools administer aptitude tests for kindergarten entry, though private schools may consider attributes that charter/magnet schools are not allowed to consider.

on the basis of characteristics plausibly exogenous to realized school type.

1.3.1 Methodology

To sidestep the endogeneity issues facing a model of school switching, the kindergarten selection model will consider only the first observed kindergarten enrollment, regardless of whether the child subsequently remained at the school for six days or six years. For students who repeat kindergarten, only their first enrollment from their first year of kindergarten is included.

Schools included in the model fall into four broad categories: traditional public (reference category, $j = 0$), public magnet ($j = 1$), public charter ($j = 2$), and private (religious or secular, $j = 3$). Using a multinomial logistic regression, probability of selecting school type j is

$$\pi_{ij} = \frac{\exp\{\eta_{ij}\}}{\sum_{k=1}^{J} \exp\{\eta_{ik}\}} \tag{1.1}$$

where $\eta_{ij} = \beta_0 + \alpha_j ECW_i + \gamma_j EM_i + X_i'\beta_j + \lambda_{zip}$. The vector X_i includes indicators for gender and race.

Some alternative school types are excluded. First, I exclude schools specifically targeted to special needs populations (e.g., schools for the blind) on the assumption that parents who select these schools are driven by a different set of concerns than parents who enroll their children in regular schools. Special needs schools include public and private options. Second, cyber schools (public and charter) and home schooling were excluded. Enrollment in these options for kindergarten is quite low, and it is possible that these parents were never inclined to pick a regular schooling option to begin with.

Multinomial logistic choice models assume independence from irrelevant alternatives – that is, eliminating a choice parents did not select would not cause them to change their selection. As noted previously, it is possible that parents observed in the data are making a temporary choice, that is, they are waiting for a spot in their preferred school to open (Engberg et al. 2014). Suppose a child is temporarily enrolled in traditional public kindergarten while their parent tries to get them a spot in a charter school, but, if that charter school did not exist, the parent would have already enrolled them in a private school. The cautious distinction to make here is that a model of selection for first kindergarten enrollment is not necessarily a model of long-term school enrollment choices.

	(1)			(2)			(3)		
	Magnet	Charter	Private	Magnet	Charter	Private	Magnet	Charter	Private
Early Child Welfare	0.45***	0.50***	0.26***	0.42***	0.48***	0.27***	-	-	-
	(0.04)	(0.06)	(0.05)	(0.04)	(0.06)	(0.05)			
Early Medicaid	0.82*	0.92	0.26***	0.89	0.94	0.34***	0.81**	0.88	0.30***
	(0.09)	(0.09)	(0.0)	(0.07)	(0.08)	(0.03)	(0.07)	(0.08)	(0.03)
Female	0.98	1.07	1.04	0.98	1.04	1.04	0.98	1.04	1.05
	(0.05)	(0.06)	(0.07)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)	(0.07)
Black	1.47	0.89	0.15***	0.78	0.57***	0.09***	0.77	0.56***	0.09***
	(0.36)	(0.18)	(0.04)	(0.17)	(0.13)	(0.03)	(0.16)	(0.12)	(0.03)
Other minority	0.82	0.51***	0.31***	0.68**	0.43***	0.24***	0.68**	0.44***	0.24***
	(0.14)	(0.11)	(0.04)	(0.13)	(0.09)	(0.02)	(0.13)	(0.09)	(0.02)
Observations	8,379			8,379			8,379		
Year FEs	Y			Y			Y		
Zip code FEs	N			Y			Y		

Standard errors clustered at the zip code level. *** p<0.01, ** p<0.05, * p<0.1

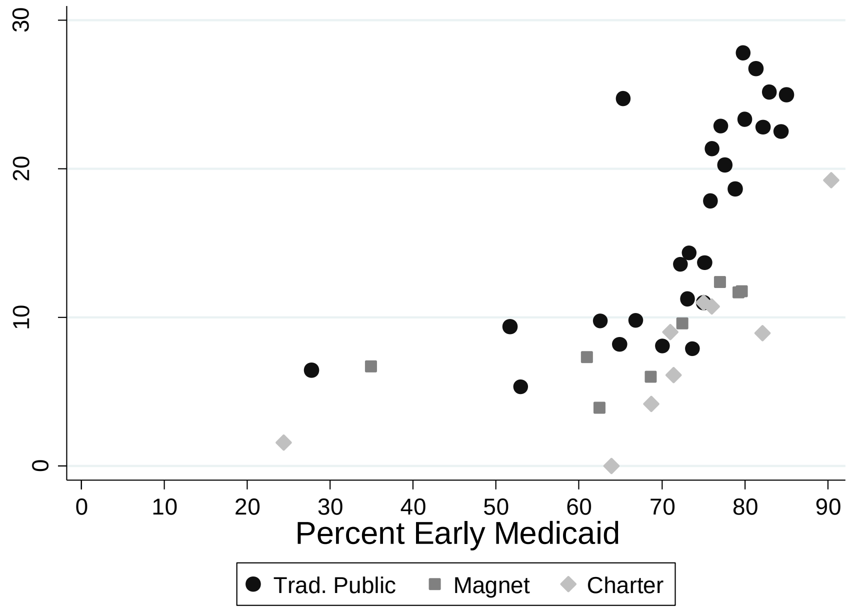
Coefficients are displayed as relative risk ratios: the null hypothesis is coeff.=1.

Figure 1.4: Multinomial logistic regression, Probability of school type enrollment relative to traditional public schools.

1.3.2 Results

Figure 1.4 shows the selection model results. Coefficients are risk ratios – that is, if they are larger than one then an increase in the corresponding independent variable predicts increased likelihood of choosing the relevant school type over traditional public school; if it is smaller than one it predicts decreased relative likelihood. Panels 2 and 3 include residential zip code fixed effects, while Panel 1 does not. Standard errors for all specifications are clustered at the zip code level. The fact that including zip code fixed effects has little impact on the estimates for ECW and EM suggests that location – while an important factor in school selection broadly – does not explain the relationship between early human services exposure and school choice.

Early Child Welfare is associated with substantially lower probability of enrollment in magnet, charter, or private schools. In fact, besides the "other minority" indicator, it is the only variable that predicts lower take-up of choice schools in general. Early Medicaid students are slightly less likely to enroll in magnet and charter schools, though the coefficients are not statistically significant in the preferred specification (Panel 2). Unsurprisingly, EM predicts substantially lower enrollment in private schools, where most students have to pay tuition.



Note: Each point represents the weighted three-year average demographic composition of kindergarten cohorts in one of the city's schools. Schools enrolling fewer than 30 kindergartners over three years were excluded from the figure. Private schools (n=16) are excluded for visual clarity, but all except one have less than 60% Early Medicaid, and all have less than 10% Early Child Welfare.

Figure 1.5: Percent of school's kindergarten cohort linked to child welfare investigations in early childhood versus percent enrolled in Medicaid in early childhood.

Panel 3 estimates the same model as panel 2 except it excludes the Early Child Welfare variable. This exclusion causes the relative probability estimates of Early Medicaid students enrolling in choice schools to decline slightly, but their confidence intervals still overlap with the estimates in Panel 2. Thus, a simple low-income proxy cannot identify the subset of disadvantaged students unlikely to access choice schools.

Selection patterns of ECW students show up in the aggregate at the school level. Figure 1.5 plots the three-year weighted average kindergarten demographic profile for each traditional public, magnet, and charter school in the city. Magnet and charter schools have similar income compositions to traditional public schools, with the vast majority enrolling a disproportionate share (more than 63%) of Early Medicaid students. The ECW dimension, however, exposes substantial heterogeneity in the composition of magnet and charter schools compared to traditional public schools. Magnet and charter schools enroll far fewer ECW students in proportional terms than do

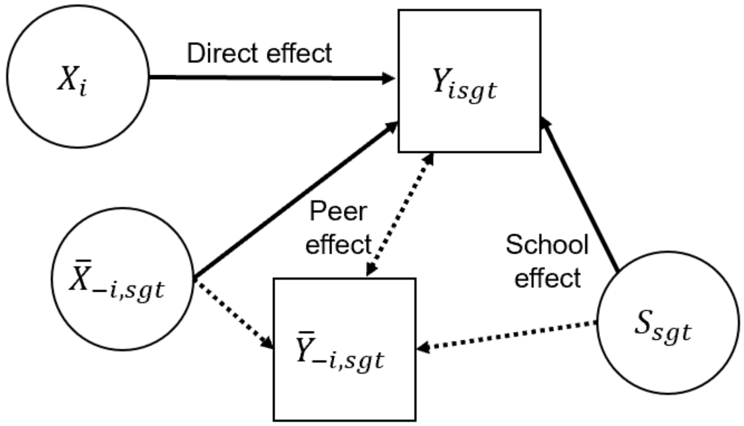


Figure 1.6: Causal pathways under exogeneity of X_i and S_{sgt}.

traditional public schools. In fact, only one choice school enrolls a disproportionate share of ECW students, while more than half of traditional public schools do. Private schools (omitted from the figure for visual clarity) universally enroll kindergarten cohorts that are less disadvantaged (on both dimensions) than the city's overall demographics.

Figure 1.5 points to the same conclusion as the selection model: a simple low-income proxy variable cannot reliably identify the schools that serve the most disadvantaged students. To my knowledge, this analysis is the first to find evidence of a systematic disparity in choice school enrollment for disadvantaged children. This disparity may be driven by disadvantaged parents choosing not to choose, and accepting a default traditional public school enrollment without exploring alternatives. I turn next to the question of how ECW students fare in schools, and whether these enrollment disparities carry implications for outcomes via peer effects.

1.4 Outcomes Model

To consider the potential pathways through which ECW and EM might impact educational outcomes, consider Figure 1.6. The goal is to understand the causes of an educational outcome, Y_{isgt}, for student i enrolled in grade g in school s during period t. The subsequent analysis models elementary school absence and suspension outcomes. The vector X_i denotes all the attributes that contribute to a student's educational outcomes over which a school has no control – including attributes of their family and home environment. Researchers cannot hope to observe all the components of X_i, but examples of X_i observed in the subsequent analysis include ECW, EM, gender and race. For each school-grade cohort in each time period, average peer attributes are constructed by calculating the average attributes of student i's cohort, leaving out student i. These average peer attributes form the vector $\bar{X}_{-i,sgt}$, and peer average outcomes (constructed analogously) are denoted $\bar{Y}_{-i,sgt}$). Finally, S_{sgt} denotes a vector of school attributes over which students have no control. These attributes may vary by grade.

There are five potential effects on outcomes in this framework. The first three are illustrated in Figure 1.6 under the assumption that X_i and S_{sgt} are exogenous. The last two effects will consider violations of those exogeneity assumptions.

1. Direct effect: student i's attributes and those of their family and home have a direct impact on their educational outcomes. Since X_i is exogenous (e.g., permanent characteristics or variables measured prior to school enrollment), there is no possibility for Y_{isgt} to, instead, cause X_i.

2. Peer effect: the average attributes of peers impact student i's outcomes. Since $\bar{X}_{-i,sgt}$ is exogenous, it cannot be influenced by Y_{isgt}. Notably, the peer effect can also operate in an endogenous fashion, in which student i's outcomes are affected by their peer's outcomes (not attributes). Hence, the graph shows pathways for both the exogenous peer effect (the direct pathway from $\bar{X}_{-i,sgt}$ to Y_{isgt}) and the endogenous peer effect (where $\bar{Y}_{-i,sgt}$ impacts and is impacted by Y_{isgt}). As indicated by the double-ended arrow between peer average outcomes and student i's outcomes, this relationship is subject to simultaneity. Thus, the subsequent estimates will identify an overall net peer effect as opposed to separating its endogenous and exogenous components.

3. School effect: school s impacts student i's outcomes directly through its policies and practices, and via the endogenous peer effect pathway by impacting peer outcomes. As long as X_i and S_{sgt} cannot impact each other, peer and direct effects can be separated from the school effect. Just as school attributes may vary by grade, so may school effects, so this can also be conceptualized as the school-grade effect.

The assumption of exogeneity of X_i and S_{sgt} is reasonable on its face – schools cannot change permanent or predetermined attributes of students, and individual students likely cannot change school policy – but comes into question when sorting and endogenous school responses come into the picture.

4. Sorting: parents select schools based, in part, on their children's attributes (which include parents' own attributes). Perhaps parents want their children to attend school with neighbors or family members, or school attributes appeal to certain types of parents/students. This leads to a correlation between variables determined at the school-grade level (S_{sgt}, $\bar{X}_{-i,sgt}$, and $\bar{Y}_{-i,sgt}$) and X_i, thus conflating peer, school, and direct effects on outcomes by breaking exogeneity of X_i.

5. School response effect: schools change their policies in anticipation of or in response to the composition and/or outcomes of their student body. This breaks the exogeneity of S_{sgt} to both student attributes and outcomes. Ultimately, time-varying school response effects coincide with peer effects, and longer-lasting school response effects coincide with sorting patterns and school effects. This simultaneity problem prevents separate identification of these effects.

It is not reasonable to assume away any of these effects. In fact, one of the motivations for investigating the effect of Early Child Welfare on outcomes is the evidence that it predicts differential sorting into schools. Thus, the empirical strategy will control for sorting, while providing estimates of direct and peer effects, albeit with some portion of the school response effect folded in.

A note on teacher effects: the choice not to give teachers their own pathway in Figure 1.6 is more an anticipation of data limitations than a statement of theory. However, the data includes only elementary grades, so the school effect can be thought to include the average effect of a given school's teachers on student outcomes over time. It is likely that teachers comprise a substantial component of the school effect on student outcomes, but, ultimately, their impact will be lumped in with that of the school. It is also worth noting that teachers generally do not have direct control over the outcomes being considered: they cannot force parents to bring their children to school, and suspension decisions are typically made by school-level administrators using guidelines set at the district level.

Before moving into the estimation strategy, it is worthwhile to align the effects discussed above with the standard enumeration of challenges to identifying peer effects: simultaneity/the reflection problem, sorting, and correlated shocks (Manski 1993; Moffitt 2001). First, simultaneity is avoided by measuring the peer effect of characteristics that are permanent or determined prior to school enrollment (in reference to the Figure 1.6, peer effects will be estimated as the net impact of $\bar{X}_{-i,sgt}$ on Y_{isgt}, rather than the effect of $\bar{Y}_{-i,sgt}$ on Y_{isgt}). The sorting problem is clearly central to this analysis and will be addressed in the identification strategy. Finally, correlated shocks would impede analysis if changes in peer demographics correlated with other unobserved shocks. While the method used will rule out a number of unobserved shocks, I will additionally run falsification tests to examine this possibility.

1.4.1 Identification & Methodology

Conceptually, peer effects seem more complicated than direct effects, but their interpretation is quite literal, which makes explaining their identification easier. Thus, I will first describe how peer effects of Early Child Welfare and Early Medicaid are identified, and then explain the identification of their direct effects.

The outcomes model seeks to estimate the impact of peer percent ECW and EM on student outcomes. The approach taken here is similar to Carrell & Hoekstra (2010) in which an extensive set of fixed effects controls for sorting. First, consider a naïve linear specification:

$$Y_{isgt} = \beta_0 + \beta_1 ECW_i + \beta_2 E\bar{C}W_{-i,sgt} + \beta_3 EM_i + \beta_4 E\bar{M}_{-i,sgt} + X_i'\alpha + \bar{X}'_{-i,sgt}\gamma$$

$$+\delta_{zip} + S'_{sgt}\eta + \theta_{gt} + \epsilon_{isgt}$$

where X_i includes permanent or pre-determined student attributes (excluding ECW and EM) for which I can also construct peer averages. Residential zip code fixed effects (δ_{zip}) are included to control for the average effect of neighborhoods on outcomes, as well as underlying sorting attributable to neighborhood (and not school) attributes. School attributes (S_{sgt}) control for school characteristics that might be correlated with both outcomes and sorting: for example, suppose smaller average class size both improved attendance and attracted parents with certain underlying attributes. Grade-year fixed effects capture the fact that students tend to miss fewer days of school but receive more suspensions as they ascend grades (see Figure 1.3). They also subsume district- or state-level policy changes and annual changes to administrative systems or variable definitions that may systematically impact outcome measures.

Of course, it is difficult to make the case that school attributes (S_{sgt}) and neighborhood fixed effects (δ_{zip}) entirely capture school effects and sorting. For one thing, the vector of school attributes would have to be quite extensive in order to capture all the features parents value. Moreover, as the selection model showed, sorting into schools is not explained wholly by neighborhoods.

Thus, the naïve model needs to be amended to include school fixed effects, which can capture the impact schools have on outcomes due to factors unobserved to the econometrician. School fixed effects also capture the average impact on outcomes of unobserved variables that correlate with sorting. Due to the concern that sorting and school effects might change over time or across grades, school-by-year (α_{st}) and school-by-grade (λ_{sg}) fixed effects will be used in the preferred specification.

$$Y_{isgt} = \beta_0 + \beta_1 ECW_i + \beta_2 E\bar{C}W_{-i,sgt} + \beta_3 EM_i + \beta_4 E\bar{M}_{-i,sgt} + X_i'\alpha + \bar{X}'_{-i,sgt}\gamma$$

$$+\delta_{zip} + S'_{sgt}\eta + \theta_{gt} + \alpha_{st} + \lambda_{sg} + \epsilon_{isgt}$$

The identifying variation in peer ECW, therefore, comes from variation in peer ECW that is not attributable to trends for grades over time (e.g., this year the district has more ECW kindergartners than usual), trends for schools over time (e.g., this year this school is attracting fewer non-ECW students), of permanent differences within schools across grades (e.g., non-ECW students tend to switch out of this school before third grade). What does that leave? Variation within a school-grade over time, demeaned from district- and school-level trends. If this variation is as-good-as-random, then β_2 has a causal interpretation as the peer effect of ECW on Y_{isgt}, with one important caveat: the time-varying school-grade response effect.

Without a detailed accounting of a school's responses to its student body composition or an instrument, the time-varying school response effect cannot be separately identified from the peer effect. This is because the portion of the school response effect that varies within school-grades over time (e.g., an individual teacher's response to their classroom composition) is determined simultaneously with the peer effect. However, for our purposes, this is fine: our peer effect estimate will include the ways in which school-grades react to variation in peer group composition, which certainly belongs in our accounting of how concentrating at-risk students impacts schools.

Ideally, school response effects could also be separated from school direct effects. It would be nice to know if the concentration of various subgroups in schools causes schools to change their policies or practices. However, these shifts are simultaneous with sorting, so the fixed effects will subsume them as they do the school direct effects.

A potential threat to identification of peer effects would be a parental response to peer groups (or contemporaneous sorting), in which parents are more likely to remove their children from a cohort during the year once they realize that cohort is particularly disadvantaged. Appendix C estimates the probability that students exit during the year using the same framework employed for the outcomes analyses. No evidence of contemporaneous sorting is found.

The external validity of the peer effects estimate depends not only on a host of contextual features of the educational system (e.g., school district suspension policies), but also on the manner in which agencies responsible for handling child welfare investigations record referrals. But here is where the scope for concerned parties to take initiative offers a benefit: the set of such concerned individuals and the environments in which they may observe child maltreatment are broad. Plus, mandated reporter laws are similar across states. Similarly, while Medicaid programs vary across states, the population identified by early involvement in Medicaid is likely broadly consistent across states.

Turning now to estimation of the direct effect: with sorting controlled for, there is no possible feedback loop through which school outcomes could influence ECW or EM – they are determined prior to school enrollment. Thus, β_1 identifies the average disparity in outcomes experienced by Early Child Welfare students, irrespective of which school they attend, the demographics of their classmates, and their own demographic characteristics (including EM, gender, race, and residential zip code).

Interpreting β_1 as the causal effect of Early Child Welfare runs into two challenges: Early Child Welfare is difficult to conceive as a treatment variable, and, as with most investigations into the social determinants of outcomes, I cannot rule out omitted variable bias. On the former matter, recall that β_1 is not the causal effect of a child welfare investigation; it is the causal effect of being in the sort of family or environment that prompts someone to report a young child to child welfare. ECW implies, by definition, that someone observed something about a child's welfare that the econometrician does not. One must be careful not to assume that the unobserved component is necessarily abuse or neglect, just as it should not be assumed that the child welfare investigation resolved (or did not resolve) the underlying issue. Potentially, ECW is indicative of a lack of parental savvy or resources alongside its most disturbing potential explanations. Still, ECW captures student attributes that are outside of schools' control, making the estimation of its effect informative to educators, case workers, and policymakers.

Returning to the possibility of omitted variable bias, of particular concern are variables that might impact both ECW and outcomes. For example, if young parents are more likely to be referred to child welfare, but are also (and independently) less concerned with regular school attendance, then some component of β_1 might reflect the effect of having a young parent, conflated with the effect of ECW. Reassuringly, the effects shown in the subsequent section are quite stable to the inclusion of controls – like race and zip code – that are also likely correlated with omitted variables of interest, suggesting that omitted variables may not loom large.

Similarly, the estimate of the direct effect of EM (β_3) requires a nuanced interpretation. It should not be interpreted as the causal effect of Medicaid, but, rather, the causal effect on educational outcomes of being the sort of child who gets enrolled in Medicaid at young ages. Ideally, there would be better measures of family income and resources available to researchers, but the direct effect of EM offers a robust and useful benchmark for the direct effect of ECW.

High take-up rates limit concerns of selection into Medicaid, but it is worth noting that β_3 could be biased by differential selection into Medicaid. If higher-income eligible families are less likely to enroll then β_3 could be biased upward; if lower-income families are less likely to enroll then the bias would more likely be downward. The intent of this paper, however, is to use EM to benchmark the effect of ECW, which, as an involuntary program, does not suffer from selection concerns.

While the absence rate outcome is estimated using linear regression, since the suspension outcome is binary, I use logistic regression with the same set of covariates and fixed effects as in the absence model.[5] Logistic regression with a large number of fixed effects can suffer from the incidental parameters problem, in which the model's estimates are not reliable. The stability of the estimates as fixed effects are added (as will be shown in Figure 1.8) suggests the incidental parameters problem may not be an issue here, likely because school-grade cohorts are sufficiently large so the model has enough data to estimate the parameters. Since the absence and suspension outcomes can be reformulated as count variables, negative binomial regression offers some potential advantages, so both the absence and suspension models are run with negative binomial specifications in Appendix A, but the results do not appreciably change.

1.4.2 Results

Figure 1.7 shows the absence model results, with each column adding more fine-grained fixed effects. Column 2 shows the results from the naïve specification, while Column 4 shows the specification that removes the effects of sorting. The direct effect of ECW on attendance is strikingly robust across specifications – absence rates for ECW students are approximately 1.5 percentage points higher than non-ECW students. This translates to a 25% increase over the mean absence rate. The direct effect of EM is also significant and robust across specifications, but is roughly half the magnitude of the ECW effect.

The direct effects of ECW and EM are not additive. Specifically, when an interaction term between the two of them is included in the regression, I cannot reject the null hypothesis that the

[5] A linear probability model is not appropriate here as suspension probabilities are fairly low, so the linear model yields a large number of negative predicted probabilities.

	(1)	(2)	(3)	(4)	(5)
Early Child Welfare	1.618***	1.624***	1.495***	1.499***	-
	(0.101)	(0.101)	(0.100)	(0.101)	
Peer	0.049***	0.066***	0.004	0.005	-
	(0.007)	(0.007)	(0.007)	(0.007)	
Early Medicaid	0.962***	0.885***	0.858***	0.863***	1.034***
	(0.061)	(0.063)	(0.064)	(0.064)	(0.065)
Peer	0.031***	0.018***	0.006	0.009	0.009
	(0.004)	(0.005)	(0.006)	(0.005)	(0.005)
Female		0.027	0.030	0.025	0.018
		(0.059)	(0.058)	(0.058)	(0.059)
Black		-0.338***	-0.331***	-0.333***	-0.318***
		(0.084)	(0.086)	(0.085)	(0.086)
Other minority		-0.288***	-0.291***	-0.291***	-0.322***
		(0.086)	(0.086)	(0.086)	(0.087)
Cohort size		0.004	-0.005	0.003	0.004
		(0.003)	(0.004)	(0.004)	(0.004)
N	43406	43406	43406	43406	43406
R-sq	0.048	0.072	0.098	0.104	0.095
Other peer controls+	N	Y	Y	Y	Y
Zip code FEs	N	Y	Y	Y	Y
Grade-by-year FEs	N	Y	Y	Y	Y
School-by-grade FEs	N	N	Y	Y	Y
School-by-year FEs	N	N	N	Y	Y

Note: Standard errors clustered at the school-grade-year level; *** p¡0.01, ** p¡0.05, * p¡0.1. **The units of the dependent variable (Absence Rate) are percentage points.**

+ "Peer" variables are calculated as the share of student i's peers - excluding student i - for whom the relevant variable equals one. Controls for gender and race peer groups are also included in specifications 2-5. None are statistically significant.

Figure 1.7: Linear regression, Absence rate.

EM direct effect for ECW students is zero. When ECW is left out of the regression (Column 5), the Early Medicaid direct effect picks up a little of the Early Child Welfare effect. This makes sense, since nearly 90% of ECW students are also EM, but the estimate remains significantly smaller than the direct effect of ECW. This again highlights how adverse circumstances beyond poverty can drive educational disparities more than poverty itself.

No peer effects are significant for absences. This is as one would expect. For young children, parents play a central role in determining whether a child attends school. It seems unlikely that a child's peers would influence the decision of a child's parents, though it is not impossible. The importance of controlling for sorting is evident in the absence specifications. In the naïve specification in Column 2, both ECW and EM are linked with significant adverse peer effects in attendance, but, once sorting into school-grade cohorts is accounted for (see Column 3), the peer effects disappear.

Figure 1.8 shows the results for the suspension model. Coefficients are expressed as odds ratios so the null hypothesis is that they equal one. ECW and EM students have substantially higher odds of receiving suspensions, and, as expected, peer effects on suspension are significant. The direct and peer effect estimates are quite similar across specifications. All else equal, the odds of an ECW student getting suspended are 57% higher than a non-ECW student; and the odds of an EM student getting suspended are 22% higher than a non-EM student. A one percentage point increase in peer percent ECW increases students' own suspension odds by 1.9%. For a one percentage point increase in EM, the peer effect impact is 1.4%. The notable persistence of gender and racial gaps in suspension across specifications will be discussed further at the end of this section.

Once again, ECW captures heterogeneity that would otherwise be lost. When it is excluded, the EM direct effect climbs slightly but not significantly so. More surprisingly, the EM peer effect does not change at all. That is, the influence of ECW students on peers is entirely missed when accounting only for poverty.

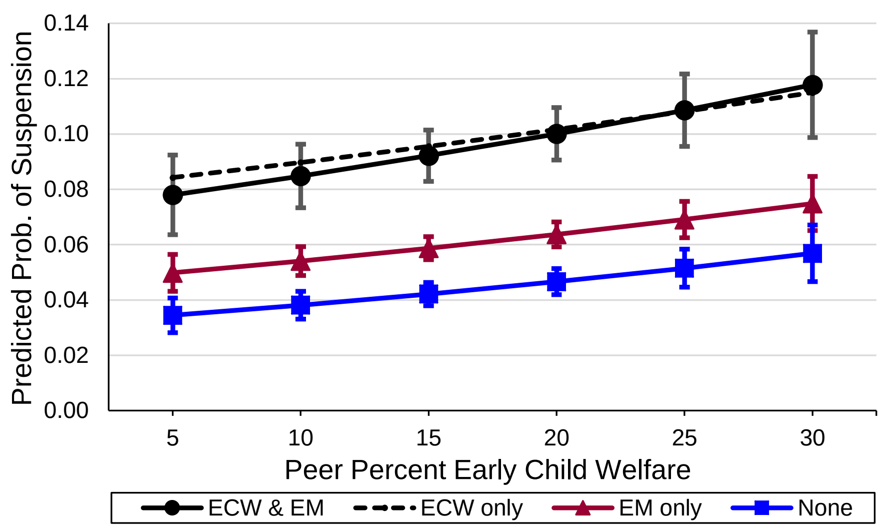
When considering the impact of sorting along the Early Child Welfare dimension, it is useful to consider how changing peer percent ECW would impact the typical student. To investigate this, I evaluate predicted suspension probabilities at the means of all the right hand side variables with the exception of peer percent ECW. I find that a typical student in a school-grade cohort that is 10% ECW (approximately the 25th percentile of the peer percent ECW distribution) has a 5.3% probability of being suspended. The same student in a school-grade cohort that is 25% ECW (approximately the 75th percentile of peer percent ECW) has a 6.9% probability of being suspended. The 95% confidence intervals of these predictions do not overlap.

	(1)	(2)	(3)	(4)	(5)
Early Child Welfare	1.613***	1.526***	1.540***	1.567***	-
	(0.059)	(0.057)	(0.060)	(0.061)	
Peer	1.035***	1.012***	1.012**	1.019***	-
	(0.005)	(0.004)	(0.005)	(0.005)	
Early Medicaid	1.385***	1.217***	1.205***	1.218***	1.280***
	(0.058)	(0.050)	(0.051)	(0.053)	(0.055)
Peer	1.011***	1.010***	1.010**	1.014***	1.014***
	(0.003)	(0.004)	(0.005)	(0.004)	(0.004)
Female		0.401***	0.388***	0.378***	0.378***
		(0.016)	(0.016)	(0.015)	(0.015)
Black		3.026***	2.940***	2.965***	2.974***
		(0.175)	(0.174)	(0.174)	(0.176)
Other minority		1.430***	1.490***	1.495***	1.483***
		(0.098)	(0.104)	(0.105)	(0.104)
Cohort size		0.999	1.007**	1.006**	1.006**
		(0.002)	(0.003)	(0.003)	(0.003)
N	43406	43379	43379	42901	42901
Other peer controls+	N	Y	Y	Y	Y
Zip code FEs	N	Y	Y	Y	Y
Grade-by-year FEs	N	Y	Y	Y	Y
School-by-grade FEs	N	N	Y	Y	Y
School-by-year FEs	N	N	N	Y	Y

Note: Standard errors clustered at the school-grade-year level; *** p¡0.01, ** p¡0.05, * p¡0.1. **Coefficients expressed as odds ratios: the null hypothesis is coeff.=1.**

+ "Peer" variables are calculated as the share of student i's peers - excluding student i - for whom the relevant variable equals one. Controls for gender and race peer groups are also included in specifications 2-5. None are statistically significant.

Figure 1.8: Logistic regression, Probability of suspension.



Note: Results from fully interacted model, where direct and peer effects can vary for each displayed subgroup. All covariates and fixed effects held to their means. Vertical bars show 95% confidence interval. Confidence intervals for the ECW only subgroup are omitted for visual clarity. They are much wider than those of the ECW & EM subgroup, but lie above those of the EM only subgroup at all points except for $x = 5$ and $x = 30$.

Figure 1.9: Predicted probability of suspension by Early Child Welfare and Medicaid exposure versus peer percent Early Child Welfare.

It is also worth investigating the possibility that different subgroups of students experience different peer effects. I run a fully interacted version of the model allowing direct and peer effects to differ based on whether students have both disadvantage indicators, just one, or neither. The predicted suspension probabilities of an average student in each of these subgroups versus peer percent ECW is plotted in Figure 1.9. The fact that each slope is positive and increasing at a similar rate shows that ECW students impact their peers regardless of their peers' own disadvantage. Moreover, the point estimates for non-low-income ECW students are right in line with those of low-income ECW students, suggesting, once again, that the challenges facing these students do not boil down simply to income.

While the previous specifications are robust to sorting across school-grades or school-years, those fixed effects subsume part of the story: parents who pick magnet schools may differ unobservably from parents who do not pick magnet schools in a manner that impacts school outcomes. Is it additionally the case that ECW or EM students in magnet schools differ unobservably from non-magnet ECW or EM students? Or, might they experience differential outcomes in magnet schools?

To inspect this, I ran the preferred specification but interacted an indicator for magnet enrollment with every covariate (including every peer measure). There is no difference in absence rates for ECW or EM students at magnet schools versus traditional public schools. Apparently, in attendance terms, ECW parents do not respond to magnet schools differently, aside from the differences attributable to school sorting. Nor is there any difference in effect of peer percent ECW on absences or suspensions in magnet schools versus traditional public schools. However, ECW students at magnet schools have significantly higher relative odds of suspension, while EM students do not. In a magnet school, the odds of an ECW student being suspended are 86% higher than a non-ECW student, while at a traditional public school, the ECW student's suspension odds are 51% higher than a non-ECW student. This suggests either that magnet schools suspend ECW students at higher rates, and/or ECW students are more likely to engage in behavior warranting suspension at magnet schools. This latter story becomes less plausible when one considers that absence rates are no different for ECW students at magnet schools, and baseline suspension rates across magnet schools are considerably more variable than those at traditional public schools. Thus, it is certainly plausible that some magnet schools suspend ECW students more aggressively, and this school response effect is being captured in the direct effect. It is important to note here that both traditional public and magnet schools fall under the same school district administration and operate under the same suspension guidelines, so this gap likely reflects implementation differences, not policy differences.

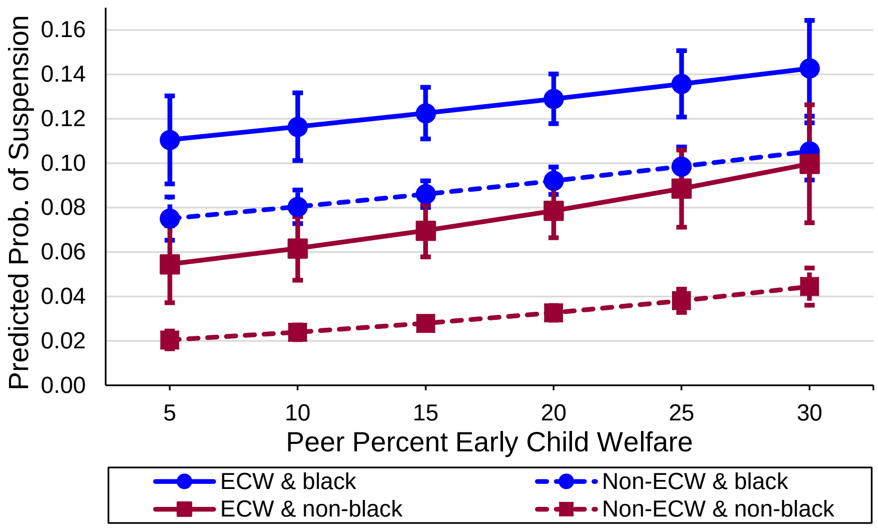
In Appendix B, the regressions are run with ECW split by whether children are linked to child welfare investigations in just one year or in multiple years prior to kindergarten. Children with multiple years of ECW fare worse than children with just one year in terms of suspension, but for absence the outcomes are approximately equivalent.

1.4.3 Results by race

Conditioning on ECW and EM, black and other minority students have absence rates that are 0.3 percentage points lower, on average, than white students (see Column 4 in Figure 1.7). The gap is reversed for suspension, in which minority students – and black students, in particular – have considerably higher suspension odds. As shown in Column 4 of Figure 1.8, suspension odds of black students are 197% larger than those of white students, while those for other minority students are 50% higher than for white students. Female students are considerably less likely to be suspended than male students. Taken together for the average student in the average cohort, this translates to baseline suspension probabilities of 2% for white girls, 6% for black girls and white boys, and more than 13% for black boys.

The ECW measure offers no explanation for this large and persistent racial disparity in suspension probability (compare Column 5 to Column 4 in Figure 1.8). Thus, it is worthwhile to explore whether direct and peer effects of Early Child Welfare persist across racial groups. In Figure 1.10 I plot suspension probabilities for black students compared to non-black students for absence and suspension. This plot is drawn from a specification of the suspension model where an indicator for black students is interacted with every individual and peer measure. As with Figure 1.9, the upward-sloping lines show that peer effects persist for ECW and non-ECW students alike, for both black and non-black students. However, a non-black student linked to a child welfare investigation in early childhood has slightly lower suspension probabilities than a black student not linked to child welfare.

As is somewhat apparent in Figure 1.10, peer effects appear proportionally larger for non-black ECW students than other subgroups. There are slight differences in the coefficient estimates of direct and peer effects of ECW by race, but all are positive and significant. Given the substantial differences in baseline suspension probabilities across subgroups, these differences can be made to seem big or small depending on how they are framed, but looking at the predicted suspension probabilities in Figure 1.10 offers a clear bottom line. ECW students have significantly higher



Note: Results from fully interacted model, where direct and peer effects can vary for each displayed subgroup. All covariates and fixed effects held to their means. Vertical bars show 95% confidence interval.

Figure 1.10: Predicted probability of suspension by Early Child Welfare exposure for black and non-black students versus peer percent Early Child Welfare.

suspension probabilities regardless of race, but these disparities operate on top of – not in place of – enormous racial disparities in suspension.

1.5 Discussion

This paper discovers a way to use administrative data to gain insight into a dimension of childhood disadvantage that has long concerned economists: disparities in parental investment in children's well-being and education that may not be attributable to disparities in family income. While more prevalent among low-income families, the effect of Early Child Welfare on educational outcomes applies whether or not the student is also low-income. By using administrative data that tracks virtually all elementary school-aged children living in a midsized city, I uncover evidence that Early Child Welfare students are less likely to enroll in charter, magnet, and private schools. In contrast, other low-income children are nearly as likely as higher-income children to enroll in these schools. Finally, Early Child Welfare students generate robust adverse peer effects in suspension that operate independently from the peer effects of low-income children.

This paper joins others in highlighting the need for better measures of childhood disadvantage than what typical school data permit. Variables more nuanced than simple poverty proxies are needed to understand the boundaries that prevent some children from accessing the best educational opportunities available to them. After all, if the costs to parents of school choice enrollment and attendance come through information and effort, not tuition dollars or transportation costs, then a proxy indicator of poverty will, at best, only partially capture the heterogeneity in family resources that may contribute to school selection and outcomes. I demonstrate this throughout my analyses by showing that when the Early Child Welfare variable is excluded, the estimates of the effect of Early Medicaid are relatively unchanged.

The findings of this paper are relevant to both child welfare and education policymakers and practitioners, including architects of school assignment systems. However, developing interventions and concrete policy recommendations on the basis of the patterns I identify in this paper requires a more detailed decomposition of the early childhood and subsequent classroom experiences of Early Child Welfare and Early Medicaid students. For example, if it can be shown that ECW is a good proxy for exposure to trauma, a case could be made for increasing school resources to support ECW students, potentially by hiring more school counselors (Reback 2010; Carrell & Hoekstra 2014). This idea aligns with evidence in the broader literature that disadvantaged students particularly benefit from increased school resources in the form of smaller class sizes (Aizer 2008; Krueger & Whitmore 1999).

However, there is also the potential for additional empirical analyses centered around policy changes and expanded data. Policy changes that are of particular interest concern school district changes in suspension practices, particularly for young children. The district under study is increasing the use of restorative practices and taking steps to reduce the number of out-of-school suspensions for young students. Future analyses can evaluate how Early Child Welfare students fare with these changes. On the expanded data front, linking information on parents' human services experiences to their children might enable an analysis of the cycles of abuse, poverty, and systemic involvement.

Appendix A: Negative binomial specification

This appendix provides an updated version of the educational outcomes models using negative binomial regressions, in lieu of linear and logistic regressions. Instead of an absence rate, the attendance outcome is the count of absent days (excused or unexcused). Instead of an indicator for whether a student is suspended during a given year, the suspension outcome is the count of suspension incidents. To account for varying lengths of enrollment – which prompt variation in the number of opportunities to be absent or suspended – the exposure variable in both models is set equal to the number of days a student is enrolled in a given school.

There are two key advantages of using the negative binomial approach. First is the ability to adjust for exposure, enabling inclusion of students enrolled for short periods of time in the model. The second advantage pertains specifically to the suspension model: including fixed effects as indicator variables in a negative binomial regression may avoid the incidental parameters problem (Allison & Waterman 2002).[6] A drawback of this approach, also noted by Allison & Waterman, is that the standard errors may be too small.

The results of the negative binomial specification (Figures 1.11 and 1.12) do not qualitatively alter the results from section IV. Specifically, an ECW student has absence incidence rates that are 23% higher than a non-ECW student, while an EM student has rates that are 18% higher than a non-EM student. In terms of suspension, an ECW student has rates that are 68% higher than a

[6]Since the dependent variables are over-dispersed – but not zero-inflated – I opt for a negative binomial specification, but Poisson regression is also consistent and does not suffer from the incidental parameters problem. The Poisson estimates are quite similar to those of the negative binomial.

	(1)	(2)	(3)	(4)	(5)
Early Child Welfare	1.264***	1.258***	1.228***	1.227***	-
	(0.017)	(0.017)	(0.016)	(0.016)	
Peer	1.008***	1.011***	1.000	1.000	-
	(0.001)	(0.001)	(0.001)	(0.001)	
Early Medicaid	1.202***	1.181***	1.178***	1.179***	1.207***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Peer	1.006***	1.003***	1.001	1.002*	1.002*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female		1.007	1.006	1.005	1.004
		(0.010)	(0.010)	(0.009)	(0.010)
Black		0.953***	0.954***	0.954***	0.958***
		(0.013)	(0.013)	(0.013)	(0.013)
Other minority		0.950***	0.947***	0.948***	0.945***
		(0.014)	(0.013)	(0.013)	(0.013)
Cohort size		1.000	0.999	1.001	1.001
		(0.000)	(0.001)	(0.001)	(0.001)
N	43416	43416	43416	43416	43416
Log pseudo-likelihood	-145133.2	-144419.7	-143730.1	-143585.8	-143764.7
Other peer controls$^+$	N	Y	Y	Y	Y
Zip code FEs	N	Y	Y	Y	Y
Grade-by-year FEs	N	Y	Y	Y	Y
School-by-grade FEs	N	N	Y	Y	Y
School-by-year FEs	N	N	N	Y	Y

Note: Standard errors clustered at the school-grade-year level; *** p¡0.01, ** p¡0.05, * p¡0.1.

$^+$ ”Peer” variables are calculated as the share of student i’s peers - excluding student i - for whom the relevant variable equals one. Controls for gender and race peer groups are also included in specifications 2-5. None are statistically significant.

Figure 1.11: Negative binomial regression, Count of absences.

	(1)	(2)	(3)	(4)	(5)
Early Child Welfare	1.869***	1.710***	1.672***	1.680***	-
	(0.081)	(0.074)	(0.069)	(0.067)	
Peer	1.042***	1.020***	1.014**	1.017***	-
	(0.005)	(0.005)	(0.006)	(0.005)	
Early Medicaid	1.440***	1.257***	1.236***	1.238***	1.306***
	(0.066)	(0.058)	(0.055)	(0.056)	(0.059)
Peer	1.012***	1.011***	1.014***	1.013***	1.013***
	(0.003)	(0.004)	(0.005)	(0.004)	(0.004)
Female		0.400***	0.384***	0.378***	0.375***
		(0.017)	(0.016)	(0.016)	(0.016)
Black		3.118***	3.052***	3.065***	3.103***
		(0.207)	(0.194)	(0.192)	(0.197)
Other minority		1.499***	1.548***	1.566***	1.581***
		(0.123)	(0.120)	(0.122)	(0.125)
Cohort size		0.999	1.007**	1.005	1.005*
		(0.002)	(0.003)	(0.003)	(0.003)
N	43416	43416	43416	43416	43416
Log pseudo-likelihood	-21312.6	-20099.9	-19507.4	-19217.0	-19299.2
Other peer controls+	N	Y	Y	Y	Y
Zip code FEs	N	Y	Y	Y	Y
Grade-by-year FEs	N	Y	Y	Y	Y
School-by-grade FEs	N	N	Y	Y	Y
School-by-year FEs	N	N	N	Y	Y

Note: Standard errors clustered at the school-grade-year level; *** p¡0.01, ** p¡0.05, * p¡0.1.

+ ”Peer” variables are calculated as the share of student i’s peers - excluding student i - for whom the relevant variable equals one. Controls for gender and race peer groups are also included in specifications 2-5. None are statistically significant.

Figure 1.12: Negative binomial regression, Count of suspension incidents.

non-ECW student, while the disparity for EM students is 24%. The increased disparity for ECW students relative to the main specification indicates that ECW students are more likely to receive multiple suspensions relative to other students.

One slight difference from the main specification is that Figure 1.11 shows significant – but very small – peer effects for Early Medicaid students in attendance (at $p < .1$). This is not surprising since the EM peer effects in the main attendance specification are of similar magnitude, and nearly as precise.

Appendix B: Heterogeneity in exposure to disadvantage

Early Child Welfare and Early Medicaid encompass a range of adverse early life experiences. The data, however, show only whether a child was linked to an investigation or enrolled in Medicaid within a given year (defined to correspond with the school calendar as August of one year through July of the next). This enables two potential variants on the disadvantage variables: a measure of the number of years of involvement a child has prior to kindergarten, and a measure of approximate age at first or last involvement prior to kindergarten. Given the limited number of years a child receives services prior to kindergarten (in which most children enroll at age 5), these potential variants are largely redundant to each other. For the sake of simplifying the following analysis, I will focus on the number of years prior to kindergarten a child is linked to child welfare or Medicaid.

For Early Child Welfare, roughly equal proportions of children are linked to child welfare for one, two, or three or more years prior to kindergarten. Without overextending the interpretation of the data, there is some more nuance about early childhood experiences that can be gleaned by comparing children linked to an investigation during one year to those linked for multiple years. Suppose a referral was made because the reporter was biased or misunderstood a child's circumstances, not because of clear evidence of a threat to a child's well-being. When the child welfare agency investigates, if they do not find any evidence of such a threat, the case can be resolved quickly. The likelihood of another investigation is low. For children linked in multiple years, on the other hand, they may be repeatedly or continually exposed to dangerous circumstances. Thus, even though this is not a precise measure of severity, children linked to an investigation during just one year may have substantively better early childhood experiences than those linked in multiple years.

In contrast, there is little variation in the number of years children are enrolled in Medicaid.

Among kindergartners in the outcomes dataset, 30% are never enrolled in Medicaid, and 55% are enrolled for 5 or 6 years prior to kindergarten (i.e., essentially enrolled from birth). Only 16% show some sort of heterogeneity, enrolling in Medicaid between one and four years. Children may have inconsistent Medicaid enrollment due to changes in household income, migration into or out of the county, or parents who choose to stop participating in Medicaid.

In the analysis that follows, in place of ECW, I use indicators for one year or multiple years of Early Child Welfare,[7] maintaining the reference category as children with no involvement. Additionally, in place of peer percent ECW, I use the percentage of peers with one year ECW and the percentage with multiple years ECW. Additionally, I use two "doses" of Early Medicaid in place of the EM indicator: 1-4 years, or 5+ years, along with peer measures for both of these subgroups.

The results in Figure 1.13 indicate that the effect attributed to ECW is stronger for students with multiple years of involvement than for students linked to an investigation during just one year. This finding is consistent with the intuition laid out above that postulates that children facing more difficult circumstances are more likely to have multiple years of ECW. However, students linked for just one year still have significantly worse outcomes than non-ECW students. This is consistent with findings of Ryan et al. (2018) and Fantuzzo et al. (2011) who show than unsubstantiated child welfare cases are still predictive of adverse outcomes for children. The negative effect of unsubstantiated or brief interactions with child welfare could be attributable to the negative effect of system involvement, proper referrals that could not be verified, or referrals resulting from bias, which impacts the family beyond the child welfare interaction.

The Early Medicaid estimates also yield an interesting nuance: children enrolled in Medicaid for 1-4 years in early childhood have outcomes that are no worse than children never enrolled in Medicaid. This result in line with the findings of Michelmore & Dynarski (2017) who find that students continually eligible for free lunch fare significantly worse than those who are occasionally eligible.

In terms of peer effects, the subgroups with more involvement show stronger peer effects – including significant adverse peer effects in attendance from students with 5 or more years of Early Medicaid enrollment. However, peers with limited involvement in ECW and EM still show adverse – if less precise – peer effect estimates.

Unlike the main specification where the effects of disadvantage were not additive, interaction

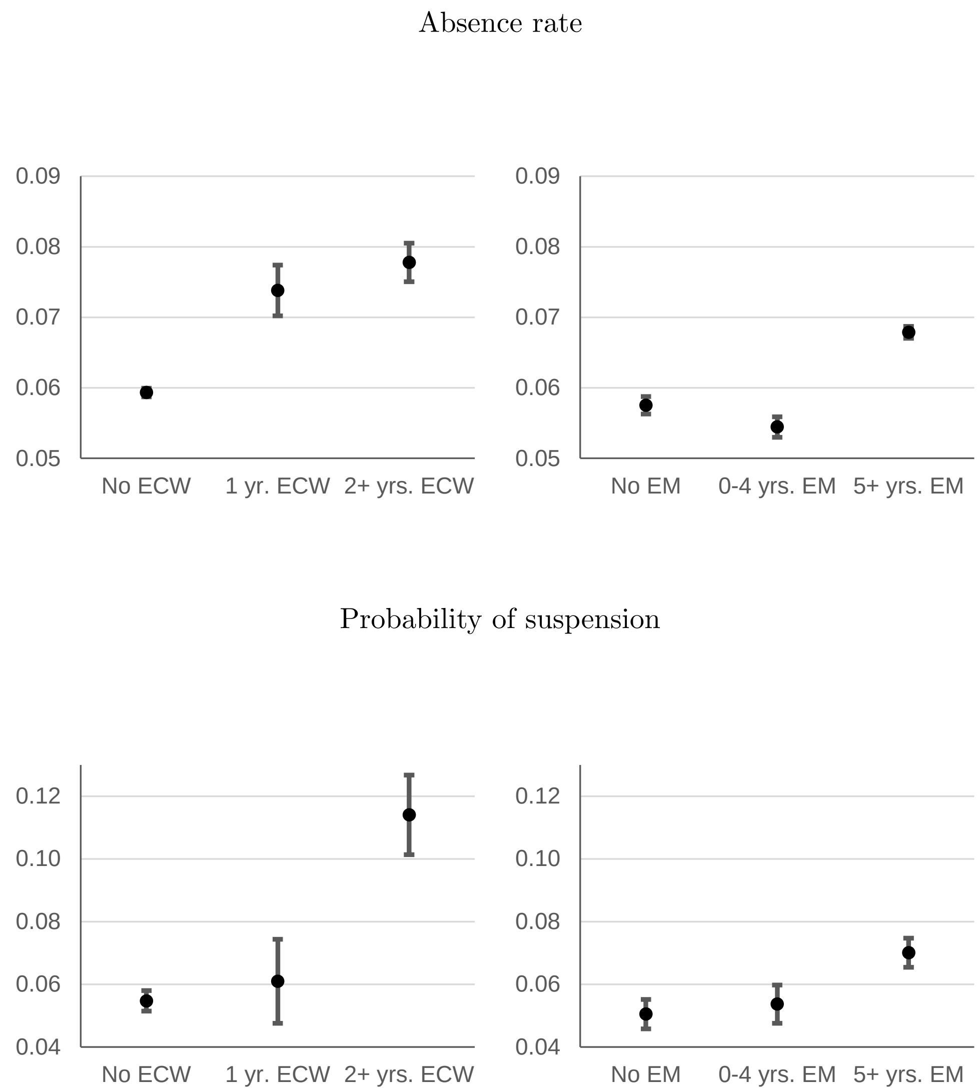
[7] This analysis was initially run with ECW doses of 2 or 3+ separated, but the results are statistically equivalent for those two subgroups. Combining them permits more precise estimates.

	(1) Absence rate *(percentage pts.)*	(2) Suspension *(odds ratio)*
1 year ECW	1.017***	1.251***
	(0.148)	(0.069)
Peer	-0.010	1.012
	(0.010)	(0.008)
2+ years ECW	1.541***	1.694***
	(0.120)	(0.044)
Peer	0.011	1.021***
	(0.008)	(0.005)
1-4 year EM	-0.029	0.993
	(0.080)	(0.062)
Peer	0.004	1.010*
	(0.007)	(0.005)
5+ years EM	1.184***	1.279***
	(0.073)	(0.045)
Peer	0.012**	1.016***
	(0.006)	(0.004)
Female	0.019	0.377***
	(0.060)	(0.041)
Black	-0.457***	2.904***
	(0.088)	(0.058)
Other minority	-0.295***	1.495***
	(0.087)	(0.070)
Cohort size	0.004	1.005**
	(0.004)	(0.003)
N	43416	42911
R-sq	0.106	
Other peer controls+	Y	Y
Zip code FEs	Y	Y
Grade-by-year FEs	Y	Y
School-by-grade FEs	Y	Y
School-by-year FEs	Y	Y

Note: Standard errors clustered at the school-grade-year level; *** p¡0.01, ** p¡0.05, * p¡0.1.

+ "Peer" variables are calculated as the share of student i's peers - excluding student i - for whom the relevant variable equals one. Controls for gender and race peer groups are also included in specifications 2-5. None are statistically significant.

Figure 1.13: Outcomes regressions with heterogeneous disadvantage.



Note: Results from fully interacted model, where direct and peer effects across all covariates can vary for each displayed subgroup. All other covariates and fixed effects are held to their means. Vertical bars show 95% confidence interval.

Figure 1.14: Predicted outcomes by heterogeneous disadvantage subgroups.

terms tell a different story in this model. This is best understood by investigating the predicted absence and suspension rates for students by disadvantage subgroup, as shown in Figures 1.14. These predictions reflect a version of the model where the effects of all covariates are allowed to vary for each subgroup. While the results in Figure 1.13 imply that students with two or more years of ECW have higher absence rates than students with just one year of involvement, the difference between the predicted attendance rates is quite small. Even one year of ECW is predictive of high absence rates. The same story is not true for suspension: predicted suspension rates for students with multiple years of ECW are nearly double those of students with 0-1 years of ECW involvement.

With respect with Early Medicaid, the predicted outcomes are consistent with the regression results in B1, with the notable observation that children with 0-4 years EM actually have lower predicted absence rates than those with no Medicaid involvement. This could reflect selection into Medicaid among households near the income cut-off, or increased efforts by schools to encourage attendance among EM students (to the extent they can identify these students, possibly through free/reduced lunch eligibility). While breaking the ECW variable out by dosage does not yield a clear story across these outcomes, the patterns for EM make the case that continual Medicaid involvement is a more useful predictor than any Medicaid involvement.

Appendix C: Probability of mid-year school exit

One potential threat to the identification of peer effects would be selective exit, where non-disadvantaged students are more likely to withdraw from a given school-grade during the school year when their peer group is more disadvantaged. This would lead to a change in peer group composition over time that might be conflated with a peer effect, since disadvantaged students with worse outcomes would be more likely to remain in the cohort.

The data enable a direct test of this possibility. I designate a student as exiting mid-year if they (a) are enrolled within the first 10 days of the school year, and (b) exit before the last month of the school year. Recall, the outcomes analysis is limited to traditional public and magnet schools. Early exits may include expulsions, which should be rare in elementary grades, but cannot be reliably separated from voluntary exits in the data.

Using the same logistic regression approach used in the suspension outcomes model in the main text, I find that, within a given school-grade, having more disadvantaged peers does not make

	(1)	(2)	(3)	(4)
Early Child Welfare	1.536***	1.536***	1.441***	1.447***
	(0.073)	(0.074)	(0.071)	(0.071)
Peer	1.043***	1.042***	1.003	1.003
	(0.003)	(0.004)	(0.004)	(0.003)
Early Medicaid	0.826***	0.796***	0.792***	0.792***
	(0.035)	(0.034)	(0.034)	(0.035)
Peer	0.995**	0.994**	0.998	0.998
	(0.003)	(0.003)	(0.003)	(0.003)
Female		0.920**	0.923**	0.924**
		(0.035)	(0.035)	(0.035)
Black		1.528***	1.517***	1.518***
		(0.092)	(0.095)	(0.096)
Other minority		1.603***	1.581***	1.572***
		(0.112)	(0.112)	(0.112)
Cohort size		0.998	1.002	1.000
		(0.001)	(0.002)	(0.002)
N	42468	42468	42468	42468
Other peer controls+	N	Y	Y	Y
Zip code FEs	N	Y	Y	Y
Grade-by-year FEs	N	Y	Y	Y
School-by-grade FEs	N	N	Y	Y
School-by-year FEs	N	N	N	Y

Note: Standard errors clustered at the school-grade-year level; *** p¡0.01, ** p¡0.05, * p¡0.1.

+ "Peer" variables are calculated as the share of student i's peers - excluding student i - for whom the relevant variable equals one. Controls for gender and race peer groups are also included in specifications 2-4. None are statistically significant.

Figure 1.15: Logistic regression, Probability of mid-year exit from school.

a student more likely to withdraw mid-year (see Figure 1.15). Thus, on average, idiosyncratic[8] changes in the percentage of peers that are Early Child Welfare and Early Medicaid do not result in the exit of non-disadvantaged students from the classroom.

However, this analysis does highlight an interesting difference among these two categories of disadvantage: Early Child Welfare students have exit probabilities 45% larger than those of other students, while Early Medicaid students are less likely to exit mid-year than other students. This gets to the heart of the potential difference between these two measures of disadvantage, in which Early Child Welfare captures a subgroup of children confronting systematically different treatment by parents and schools (in the case of expulsion) that is not exclusively attributable to poverty.

[8] Idiosyncratic in the sense that this variation is not attributable to systematic sorting into schools over time, or sorting in and out of the district over time.

Chapter 2

An Intervention to Reduce Chronic Absenteeism

2.1 Introduction

Chronic absenteeism presents a major issue for public schools. When students miss school, they miss academic instruction and other in-school services. Absences causally reduce achievement, especially for low-income students, who are much more likely to be chronically absent (Gershenson et al. 2017; Goodman 2014; Gottfried 2014; Romero & Lee 2007). Chronic absenteeism is also linked to school drop-out and adverse outcomes in adulthood, such as unemployment, criminality, and drug and alcohol abuse (Sutphen et al. 2010; Rocque et al 2017).

Teachers and school administrators face difficult decisions when they encounter students who are chronically absent or facing problems outside of school. As mandated reporters, they are legally obligated to report suspicions of abuse or neglect, including educational neglect. Yet for a student who is tiptoeing towards chronic absenteeism but whose safety at home is not in question, a child welfare referral may not be appropriate or productive. When absenteeism meets the legal standard for truancy, school districts can report families to the court system. Yet that entails a burdensome and costly process and is often treated as a last resort.

The need for a middle ground – a non-punitive way to confront absenteeism among primary school students – prompted local leadership in the anonymous county under study to develop a truancy prevention program. While programs aimed at improving school attendance generally are widespread, the truancy prevention program offers targeted interventions for students most at risk.

Program administrators meet with students and caregivers to identify the root causes of absenteeism, and provide services customized to students' needs (e.g., tutoring, resolving transportation issues, etc.). Participation is voluntary.

The program was piloted in the 2013-14 school year at two K-8 schools in an urban public school district. To be eligible for treatment, students had to be referred by teachers and then screened by program staff. The process did not involve random assignment.

This paper uses a difference-in-differences approach to conduct an intent-to-treat analysis of the truancy prevention program. This approach entails comparing the outcomes of all students in the pilot schools – not just those directly treated – before and after the program began to the outcomes of students enrolled simultaneously in similar schools that were not in the pilot. Identification of the treatment effect relies on two arguments: first, that students and their caregivers did not enter or exit treatment schools on the basis of the program; and second, that treatment and control schools would experience parallel trends in outcomes if not for the treatment. The methodology section discusses these assumptions in detail.

The study benefits from a rich set of controls that account for important student risk factors for chronic absenteeism. Data on human services utilization permit construction of low-income designations prior to treatment that are independent of schools (in contrast to free/reduced-price lunch designations, which may be endogenous). An analysis of students directly involved in the program reveals that persistently low-income students were considerably more likely to take up treatment services when offered them. This pattern enables a triple-difference approach, that not only compares student outcomes across treatment and control schools, but additionally compares — within schools — the outcomes of persistently low-income students to those who are not.

The existing evidence on truancy prevention programs is somewhat encouraging, but limited. Sutphen et al. (2010) conduct a systematic review of evaluations of U.S.-based truancy prevention programs and note the need for more experimental and quasi-experimental evaluations. The most similar paper to this one — Cabus & De Witte (2015) — uses a difference-in-differences approach to analyze the effectiveness of a truancy prevention program in Dutch secondary schools. They find a significant reduction in school dropout. However, a rigorous evaluation of the U.S.-based Communities in Schools case management approach found no program impacts, and the intervention bears several similarities to this one (Balu 2019). Notably, both of the aforementioned interventions are implemented in secondary schools populations, while this one is implemented in a K-8 setting.

In addition to providing direct services, the program under study necessarily involves parental engagement, which shows promise for reducing absenteeism. Smythe-Leistico & Page (2018) implement a text-message intervention for parents of kindergartners and find a substantial reduction in chronic absenteeism relative to a synthetic control. Avvisati et al. (2014) evaluate a field experiment meant to improve parental engagement in a "deprived" education district in France. They find a robust reduction in truancy for children whose parents participated in the intervention.

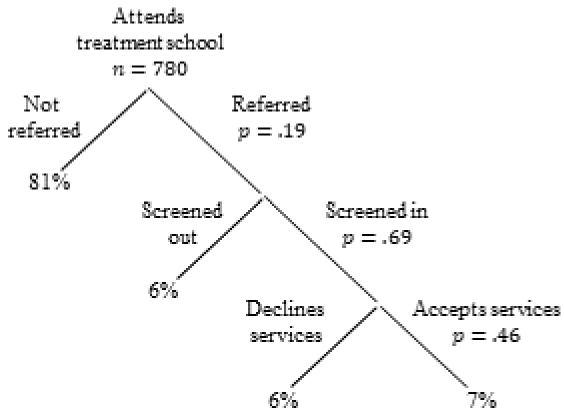
In the long run, a truancy prevention program is not simply about improving the attendance of at-risk students through direct services and parental engagement, but also curtailing the use of costly and burdensome external systems. The program under study offers teachers a way to connect students with services when they have concerns about students, but do not think child welfare or court involvement is warranted. The presence of the program may alter the way in which teachers respond to students who are missing a lot of school. Analysis of these potential outcomes, however, falls outside the scope of this paper.

This study is limited to the short-run: did the program – while underway – succeed in reducing rates of chronic absenteeism and truancy in the treatment schools? I find that the program was somewhat effective in reducing chronic absenteeism and truancy among persistently low-income students in treatment schools. This is consistent with the fact that low-income students were more likely to take up program services.

2.2 Background & Data

The program's aim is to catch a student's attendance issue early and address its root causes. Figure 2.1 summarizes program implementation within treatment schools during the 2013-14 school year. For individual students, entry into the program begins when teachers refer their names to the program administrator. 19% of students in treatment schools were referred into the program. Next, the program administrator screens student records to ensure the program is not redundant to services the student is already receiving from school staff or caseworkers. If the student is deemed eligible, the administrator reaches out to the student's caregiver for a meeting. Following the meeting, the caregiver either accepts or declines services customized for the student. In the 2013-14 pilot year, roughly one-third of referred students were screened out, one-third declined services, and one-third accepted direct services. Services are administered by program staff or contracted to an external provider.

Figure 2.1 also provides a useful breakdown of sources of selection relating to individual students



Conditional Probability	Persistently low-income	Not persistently low-income
Referred	.19	.18
Screened in	.72	.62
Accepts services	.51	.29
N direct treat	40	7

Figure 2.1: Intervention summary.

as they make their way into direct treatment. Teacher referrals are not random – inherently, they reflect teachers making a choice based on factors they observe but the econometrician cannot. They may condition their referrals on their perceptions of who could benefit most from the program's services. Screening by program administration may reflect observable administrative records – identifying children already receiving services from a caseworker, for example – but also may reflect information from other school sources including the referring teacher. Finally, choices by caregivers to accept a meeting with program staff and then take up services are also not random.

In sum, students who accept services differ from those who decline services on the basis of caregiver selection; students who are screened into the program differ from those screened out in terms of the perceived suitability of program services for the student; and teachers select students for referral on the basis of unobserved factors. A rich set of human service covariates – normally not observed in educational data or by teachers – do not reliably predict which students teachers refer to the program. Conditioning on prior year attendance records causes students who switched schools to fall out of the analysis, which represents a sizable portion of referred students. Thus, I proceed with an intent-to-treat analysis on the school level.

Program administrators began working with some students in treatment schools during the 2012-13 school year, but did not enter full implementation until the 2013-14 school year. Therefore, the pre-treatment period is the 2011-12 school year, and the treatment period is the 2013-14 school year.

2.2.1 Data

The data are compiled from three sources:

1. The program administrator maintained a log of students in the treatment schools who were referred for services. The log additionally notes whether students referred for services were offered services, and whether their caregivers accepted them.

2. Administrative enrollment and attendance records are available for all students enrolled for any length of time in treatment and control schools from school years 2010-11 through 2014-15. Enrollments are recorded as date spans, and absences are recorded by date. Attrition occurs in the data when students exit the school district, potentially through switching schools or moving. Therefore, full attendance and enrollment histories are not available for all students. These data also include students' residential zip code, as well as other standard demographic indicators like gender and race.

3. Human service records include indicators of whether students were enrolled in Medicaid or linked to a child welfare investigation each year for at least three years prior to the intervention. These records are available as long as a student resides in the city or surrounding county, so they offer reliable measures of disadvantage prior to the start of each school year, even if the student was previously not enrolled in the district. The following two covariates stemming from human service records are useful predictors of chronic absenteeism:

    - *Persistently low-income.* A student is considered persistently low-income if they are enrolled in Medicaid continuously for the prior three years.
    - *Child welfare involvment.* A student is flagged for child welfare involvement if they are linked to a child welfare investigation at any point in the prior three years.

The population of study is defined as all students enrolled in the district's neighborhood (non-magnet) K-8 schools during either 2011-12, 2013-14, or both. Students are included in the analysis if they are enrolled for at least one-half of the relevant school year. This precludes students from showing up in multiple schools during the same school year in the regression dataset.

The population varies considerably in demographic terms. Figure 2.2 summarizes characteristics of students in the treatment year, as well as baseline attendance statistics during the pre-treatment period. Sorted by the share of students who are persistently low-income, it is clear that the treatment schools – schools 1 and 2 – are two of the most disadvantaged.

School	1	2	3	4	6	5	8	7	9
Treatment	Y	Y	N	N	N	N	N	N	N
Missing 2010-11 attendance	N	N	Y	N	N	Y	Y	N	N
N*	326	180	398	250	268	413	431	277	558
Percent of students									
Persistently low-income	70	68	68	58	47	46	44	42	22
Free/reduced-price lunch	92	93	91	86	72	77	69	75	44
Child welfare involvement	28	18	25	14	10	7	9	7	6
Black	89	89	65	77	35	47	9	28	31
White	4	7	22	14	57	40	80	53	49
Special education	18	44	30	22	19	16	20	23	9
Pre-treatment (2011-12) school outcomes									
Absence rate	6.3	6.7	5.8	4.5	5.2	4.9	5.7	5.2	4.3
% of absences unexcused	68	55	61	43	34	35	39	34	40
Students chronically absent (%)	18	24	18	9	13	11	16	12	8
Students truant	74	57	66	42	38	37	40	36	33

*Includes students in grades 2-8. All statistics reflect treatment year (2013-14) data unless otherwise noted.

Figure 2.2: Demographics and absence rates by school.

The analysis focuses on two outcome variables:

- *Chronic absence.* Students are deemed chronically absent if they miss 10% or more of the school days for which they are enrolled. Detailed enrollment data permits adjustment of absence rates according to the number of days the student is officially enrolled in a given school. This measure does not distinguish between excused and unexcused absences, but suspension days are netted out of both the numerator and denominator.

- *Truancy.* Students are flagged as truant if they have three unexcused absences.

The truancy definition is based off of a state legal standard, but there are practical issues with this outcome. First, many students accumulate three unexcused absences over the course of the school year. In fact, the unexcused absence rates shown in Figure 2.2 suggest that the average student at most schools met or exceeded this truancy standard in the 2011-12 school year. Second, students can reduce their unexcused absence rate without actually improving their overall attendance, simply by submitting excuse notes from caregivers. While the program is geared towards preventing truancy, the aim is not simply to encourage caregivers to submit notes when students are absent. Finally, many students were referred after their third unexcused absence, reducing the ability of the intervention to impact on this outcome. The chronic absence standard is a better way to evaluate the practical success of the program.

2.3 Methodology

Given the fact that only 6% of students in the treatment schools received direct treatment, finding a significant treatment effect in an intent-to-treat analysis at the school level is a high bar. Treated students likely have high and highly variable absence rates, so movement in school mean attendance may not show up even if the program is effective. However, students receiving direct treatment should represent a substantial share of students at risk of chronic absenteeism. Even though treated students receive services at varying times throughout the school year, if the program is effective, it should drive down annual chronic absence rates at the treatment schools.

The reliability of treatment estimates from a difference-in-differences analysis hinges on two assumptions: no selection on the basis of treatment, and parallel trends in outcomes across treatment and control units in the absence of the treatment. The possibility of selection into or out of treatment schools on the basis of the program is unlikely – even with the soft launch in 2012 – given that only students who were truant or at risk for becoming so were offered program services. Approximately one-half of caregivers declined services. It is unlikely a student would select into or out of a school on the basis of optional services offered only to students with attendance issues.

One threat to the parallel trends assumption comes through a structural change to one of the treatment schools, in which the number of kindergarten and first grade classrooms was increased between the 2011-12 and 2013-14 school year. To avoid this issue, grades K-1 are excluded from the analysis.[1]

Aside from the aforementioned enrollment shock, since the schools are operated by the same district, any changes to district attendance policies should impact all schools. Exogenous shocks over time that impact all students – like district scheduling quirks (e.g., a one-day school week) and bad weather – lead to systematically higher or lower absence rates. These shocks should be common across schools. Of course, attendance is a common concern for teachers and school administrators, so within schools there are likely various attempts made over the course of the year to improve attendance. I cannot control for such efforts. To the extent they are generally similar (e.g., phone calls home to check on students with consecutive absences), they should wash out when comparing

[1] When the analysis is run including these grades, the estimated treatment effects have the same sign, but slightly lower magnitude and precision.

outcomes across years.

An approach that might resolve concerns of time-varying attendance policies across schools is a triple-difference specification, in which persistently low-income students are treated as the primary treatment group. The treatment effect is then estimated by comparing, across schools, the relative change in outcomes for persistently low-income students to that for non-low-income students. This approach is pursued on the basis of the treatment probabilities summarized in Figure 2.1. While students are approximately equally likely to be referred into the program regardless of income, persistently low-income children are somewhat more likely to be offered program services, and considerably more likely to accept them. If the underlying barrier to attendance for persistently low-income children is resource-based, and the program can offer some sort of resolution to this resource disparity, then higher take-up and, potentially, larger treatment effects should be anticipated for persistently low-income children.

Returning to the issue of the parallel trends assumption, issues in attendance data for some of the control schools prevents a complete analysis of pre-treatment trends. Panel A in Figure 2.3 shows patterns in chronic absenteeism among treatment and control schools, with the control schools with missing 2010 data plotted separately. Pre-treatment (pre-2012), chronic absenteeism across treatment and control schools moves together, despite clear disparities in overall rates. The fact that the trend between the two sets of control schools between 2012 and 2013 is similar is encouraging. The relatively large decline in chronic absenteeism at the treatment schools between 2012 and 2013 suggests the intervention may have had an effect.

One concerning aspect of Figure 2.3 is the apparent difference in the control group trends between 2011 and 2012. This concern is exacerbated in the truancy rate chart (Panel B) which shows the control schools with missing data in 2010 diverging from the other control schools between 2011 and 2012. A potential explanation for the issue is this: if the control schools with missing data started using the electronic attendance system in 2011, potentially they did a worse job of distinguishing unexcused absences from excused ones in their electronic records. While the DDD specification leverages a within-school comparison that may resolve some inconsistencies in attendance recording across schools, this trend issue paired with the conceptual concerns about the truancy definition suggests the chronic absenteeism outcome will yield more reliable estimates.

The difference-in-differences (DD) specification is defined as follows:

$$y_{isgt} = \alpha_0 + \sum_{s'=1}^{S} \gamma_{s'} 1\{s = s'\} + \alpha_1 \mathit{2013} + \delta \; TREAT * \mathit{2013} + \lambda_g + \eta_{zip} + X_i'\beta + \epsilon_{isgt} \tag{2.1}$$

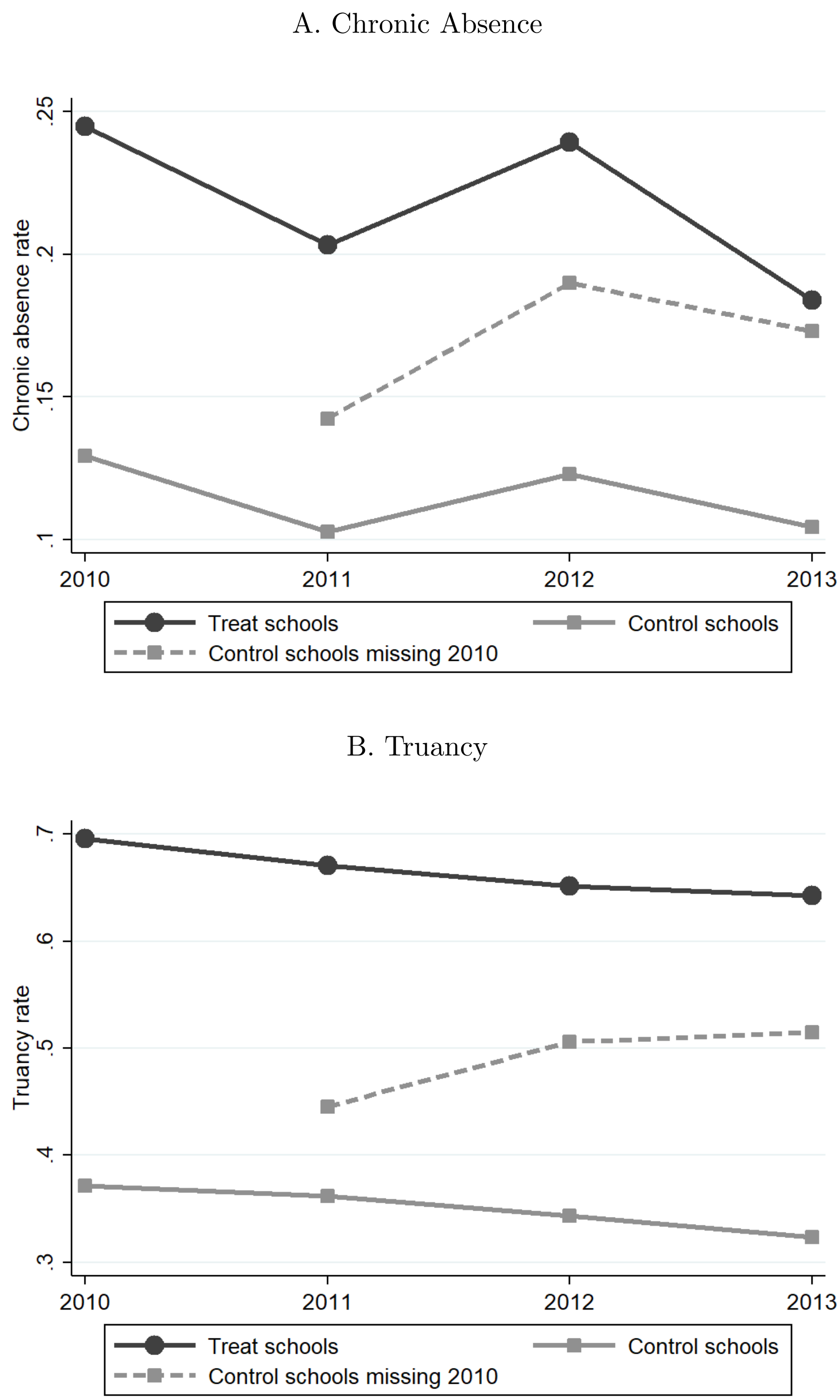


Figure 2.3: Trends in outcomes across treatment and control groups.

where

- schools are indexed $s \in \{1, \ldots, S\}$,
- $TREAT$ is an indicator for $s \in \{1, 2\}$ where schools 1 and 2 are the treatment schools,
- *2013* is an indicator for the treatment period (i.e., the 2013-14 school year),
- λ_g is a vector of grade-level fixed effects,
- η_{zip} is a vector of residential zip code fixed effects, and
- X_i are a set of individual controls (including race, gender, and human service indicators).

The treatment effect is estimated by $\hat{\delta}$.

Zip code fixed effects adjust for potential neighborhood effects in attendance. Grade-level fixed effects account for systematic variation in attendance rates across grades. Specifically, absence rates across grades are u-shaped, declining in early elementary grades before trending back up during middle school grades.

The triple-differences (DDD) specification makes an additional contrast for persistently low-income students by interacting PLI_i — an indicator for whether student i is persistently low-income[2] — with the treatment, school, and time indicators:

$$y_{isgt} = \alpha_0 + \sum_{s'=1}^{S} \gamma_{s'} 1\{s = s'\} + \sum_{s'=1}^{S} \gamma_{s'}^{L} 1\{s = s'\} * PLI_i + \alpha_1 \mathit{2013} + \alpha_2 PLI_i \tag{2.2}$$

$$+\alpha_3 \mathit{2013} * PLI_i + \alpha_4 TREAT * \mathit{2013} + \delta\ TREAT * \mathit{2013} * PLI_i + \lambda_g + \eta_{zip} + X_i'\beta + \epsilon_{isgt}.$$

The treatment effect in the DDD specification is also designated $\hat{\delta}$.

2.4 Results

Figure 2.4 shows the results for both specifications. For chronic absenteeism, the DD specification (column 1) yields a treatment effect that is negative (as anticipated), but imprecise. In the DDD specification of column 2, the treatment effect on persistently low-income students is negative and weakly significant at the 10% level. Specifically, persistently low-income students at treatment schools saw a 4 percentage point decline in their probability of being chronically absent relative to other students. Given the baseline disparity where persistently low-income students are 7pp more

[2] PLI_i is included as a covariate (in X_i) in the DD specification.

	Chronically absent		Truant	
	(1)	(2)	(3)	(4)
Persistently low-income (PLI)	0.075***	0.071***	0.142***	0.077**
	(0.013)	(0.014)	(0.023)	(0.026)
Treatment year (2013)	0.013	0.012	-0.000	0.004
	(0.015)	(0.010)	(0.024)	(0.015)
2013*PLI		-0.001		-0.012
		(0.019)		(0.034)
TREAT*2013	**-0.019**	0.008	**-0.051**	-0.003
	(0.043)	(0.044)	**(0.032)**	(0.025)
TREAT*2013*PLI		**-0.040***		**-0.068***
		(0.021)		**(0.036)**
School indicators *(subsumes TREAT)*	Y	Y	Y	Y
PLI*School indicators *(subsumes PLI*TREAT)*	N	Y	N	Y
Grade & zip code indicators	Y	Y	Y	Y
Individual controls	Y	Y	Y	Y
N	6102	6102	6102	6102
R-sq	0.071	0.075	0.183	0.188

Standard errors clustered by school. *p<0.1, **p<0.05, ***p<0.01

Figure 2.4: Linear regression, Probability of chronic absenteeism or truancy.

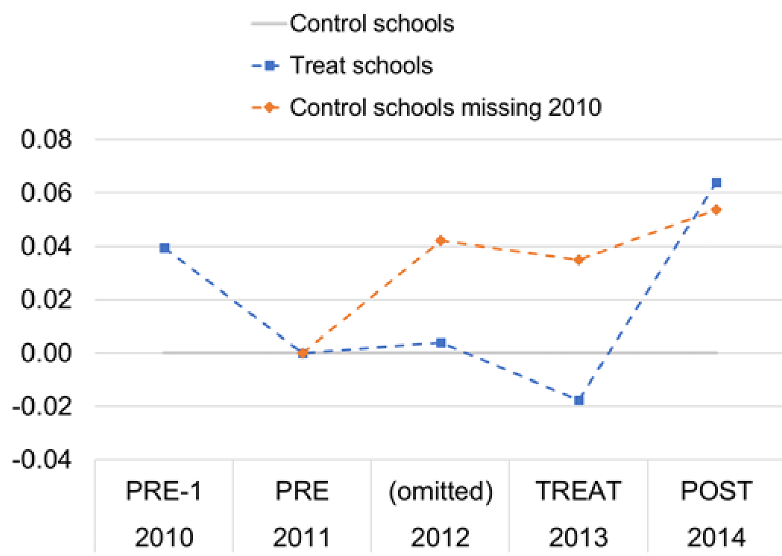


Figure 2.5: Event study, Relative probability of chronic absenteeism.

likely to be chronically absent than their counterparts, the treatment effect amounts to closing the gap by more than one-half.

The estimated treatment effects for the truancy outcome are slightly larger, but should be interpreted cautiously due to potential issues with the comparison group. The change in the baseline disparity in truancy for persistently low-income students between the DD and DDD specifications (columnn 3 and 4) from 14pp to 8pp suggests that relative truancy rates for low-income students across schools vary substantially, in contrast to the chronic absence disparities that are consistent across specifications (columns 1 and 2). This is consistent with the theory that some schools systematically varied in recording unexcused absences. If the way unexcused absences were recorded changed between the pre-treatment and treatment periods, the truancy treatment effect may be biased.

Now to revisit the issue of control schools with missing 2010-11 attendance data, for which I cannot compare pre-trends. When the DDD specification is run without the missing data control schools, the treatment effect estimate is negative but quite imprecise. In figure 2.5, I show the results from an event study based on the DDD specification. This shows how the relative chronic absenteeism rate of PLI students in treatment schools and missing data control schools evolves from the 2011-12 school year, relative to the rest of the control schools.

When treatment schools are compared to the control schools without missing data, a modest decline in chronic absenteeism during the treatment year is observed, though the difference between the two is not statistically significant. This decline notably contrasts with the growth in chronic absenteeism rates at the control schools with missing 2010 data between 2011 and 2013. That divergence in outcomes drives the magnitude of the DDD treatment effect estimate. Ultimately, the reliability of that estimate depends on whether the increase in chronic absenteeism of PLI students at the missing data control schools is real, or the result of under-reported absences in 2011-12.

2.5 Discussion

The treatment effect for persistently low-income students is encouraging, despite its imprecision. Given the costly consequences of truancy (e.g., school dropout, court involvement), effective truancy prevention programs have the potential to be quite cost effective. The voluntary nature of the program paired with the tendency of take-up to favor persistently low-income children suggests that targeting persistently low-income children may improve program efficiency. The treatment schools were disproportionately low-income, so as the program expands to schools with higher-income demographics, effective targeting may become even more important.

Intuitively, the program should work best for children whose attendance issues are driven by a lack of family resources or awareness, as opposed to caregivers who cannot be persuaded to encourage their children to achieve regular attendance regardless of resources. One additional question worth investigating concerns identifying which program services worked best and for whom. Was there a subset of children for whom outreach to caregivers with communication about the importance of attendance caused an improvement? Or is the program effect driven through the reception of direct services? Isolating these drivers would provide important guidance for efficient design and implementation of truancy prevention programs.

Chapter 3

Persistent and Wide-Ranging Differences in the Income and Racial Segregation of Children

3.1 Introduction

As school district desegregation efforts subsided and the decline in school racial segregation stalled, evidence of persistent racial segregation across school districts in metropolitan areas throughout the U.S. emerged (Rivkin 1994; Clotfelter 1999; Reardon et al. 2012). The Supreme Court case *Milliken v. Bradley* (1974) blocked efforts that reached across school district boundaries to reduce de facto school segregation. School desegregation efforts were thus unable to counter regional segregation and white flight (Reber 2005).

Racial segregation is linked with adverse education and labor market outcomes for non-white Americans, while school desegregation efforts starting in the 1970s improved outcomes (Cutler & Glaeser 1997; Reardon et al 2017; Ashenfelter et al. 2005; Guryan 2004). In light of recent research that links children's long-term outcomes to the particular neighborhood in which they grow up, residential and school segregation offers a potential mechanism through which inequalities foment (Chetty et al. 2018). Income segregation is increasingly explored alongside racial segregation in analyzing disparities. It, too, is correlated with racial achievement gaps, as well as income achievement gaps and lower levels of income mobility (Owens 2018; Chetty et al. 2014).

By revisiting geographical analyses of segregation (including Rivkin 1994, Clotfelter 1999, and

Owens 2016) but placing measures of children's income and racial segregation side-by-side, this paper uncovers clear differences in the geographical patterns of racial and income segregation among children. School district boundaries factor more into racial segregation than income segregation. Income segregation builds from variation across neighborhoods and schools within districts. These geographical differences have implications for policies that might impact segregation, whether by design or through spillover effects.

In the process of contrasting children's income and racial segregation, this paper also places neighborhood (residential) and school segregation side-by-side. Since children usually attend school near where they live – whether by policy or preference – these measures are inherently linked. Potentially, neighborhood effects on economic outcomes are as strong as school effects (Card & Rothstein 2007; Chetty et al. 2018). Yet school segregation remains unique for its central position in the debate over policies past (desegregation) and present (school assignment and choice policies). Practically, school segregation within districts can be directly augmented through policy, while residential segregation and segregation across larger geographies is subject to change through policy only indirectly.

I estimate income and racial segregation within all U.S. commuting zones using data from the year 2000, the last time the Census long-form survey asked respondents to report income. I also calculate the same segregation indices for the 1999-2000 school year using data from NCES's Common Core of Data. Taken together, these sources enable estimation of key segregation measures pertaining to children: income and racial segregation across neighborhoods, schools, and school districts. To my knowledge, this is the first paper to offer a comprehensive summary of this full set of measures across all commuting zones, nationally. I take careful steps to vet and compare inputs across data sources. Segregation measures rely on population counts or massive samples for accuracy, and the dearth of reliable, publicly-available income data on children on a national scale presents a major challenge for tracking these measures over time.

Segregation among children in 2000 was substantial. Consistent with Owens (2016), I find that children experience considerably higher levels of residential income segregation than the broader population. While residential racial segregation of children is only slightly higher than that of the adult population, residential and school racial segregation remains objectively high. On average, one-half of minority children would have to move schools and neighborhoods to create an even population spread.

Commuting zones are ideal regional definitions for several reasons: they are constructed to encompass local labor markets, they cover all 50 states and Washington, D.C., and, since they aggregate counties, they can be merged into a range of local geographies. The first point in particular makes them a good starting point for analyzing segregation; intuitively, in the absence of de jure segregation, segregation arises based on families' local residential choices – e.g., given that we work in this location, and our income and housing options are such, what is the best neighborhood/schooling options for our family? Assuming families take community attributes (e.g., housing prices, peer group composition, etc.) as given when deciding where to live, sorting likely reflects preferences for community attributes, budget constraints, and information asymmetries. Bias, too, factors in, particularly in the context of racial segregation.

Essentially, segregation indices measure sorting. As an economic theory, sorting offers predictions about how segregation might react to the geographical and structural context of a given commuting zone (Nechyba 2003; Epple & Platt 1998). For example, commuting zones vary considerably in terms of the number and typical size of school districts they contain. Inherently, when there are more options – lots of school districts with different attributes – families can more closely match their community's attributes to their preferences, and segregation across school districts, as a result, will be higher. This is consistent with the Reber (2005) finding that white flight in response to school desegregation was more severe when there were more nearby public school districts. Of course, the same logic should apply to sorting within school districts when school district concentration is high: in a commuting zone that contains just a few large school districts, school district boundaries may not create a particularly meaningful distinction between communities. Other variations in neighborhoods – like proximity to specific schools – may become more important.

To test these predictions, I regress segregation on commuting zone geographic attributes — namely, commuting zone size and the concentration of school districts within the commuting zone. I find that in regions with large school districts, district boundaries are less salient to both racial and income segregation across school *districts.* So do children in large school districts experience less segregation across schools? It depends. School racial segregation is lower with more concentrated school districts, but school income segregation is higher. The same is true for neighborhood income segregation. Families still sort by income into neighborhoods even when sorting across school districts is limited.

An alternative version of this analysis built from the inputs to segregation indices — as opposed to the indices themselves — offers a more practical summary of these geographical differences. Which boundaries explain the variation in school demographics that children experience? To investigate this, I analyze the proportion of the variation in school subgroup population shares that is attributable to sorting across various geographical levels. I find that more than half of the national variation in non-white population shares across schools is attributable to sorting across larger regions (commuting zones and counties). Only 18% of the variation in school non-white population shares occurs within school districts. In contrast, more than 40% of the variation in school low-income population shares occurs within school districts. Using additional years of school data, I show these patterns are quite persistent over time.

The patterns discussed in this paper offer important context for researchers and policymakers seeking to understand how policy might impact segregation, and vice versa. Racial and income segregation are not the same. Sorting patterns of non-white and low-income children across commuting zones, school districts, neighborhoods, and schools differ substantially and persistently. Thus, the geographical bounds of policies and their relative impact across space matter differently to their effects on the outcomes of non-white and low-income children.

3.2 Data & Geographies

All U.S. counties in 1990 were assigned to one of 738 commuting zones.[1] Commuting zones (sometimes abbreviates as “CZs”) are meant to approximate regional labor markets, and can be broken down into smaller geographic subdivisions like Census tracts (approximate neighborhoods containing about 4,000 people). School districts, which have geographic boundaries that are sometimes as large as counties or as small as towns, offer another subdivision of commuting zones. Schools, in turn, are nested into school districts.

The data is drawn from two sources: the 2000 Census[2] and the Common Core of Data (CCD). Population counts by income and race in the Census data are available at the school district and Census tract level, and broken out by relevant subgroups (e.g., households with children, children less than 18 years old). CCD offers public school enrollment counts by free and reduced-price lunch eligibility (FRPL) and race at the school level.

Figure 3.1 offers key population and geographic summary statistics of the 710 commuting zones

[1]I use 1990 commuting zones to enable comparison to Chetty et al.’s estimates. Results are unchanged when commuting zones from 2000 are used.

[2]Census tract tabulations are drawn from the American FactFinder data download tool; school district tabulations come from the National Center for Education Statistics EDGE tool.

	All commuting zones (n=710)					Limited (n=588)
	Mean	Median	Std. dev.	Min.	Max.	Mean
Households	147,882	41989	381,527	897	5,359,948	146,642
Individuals	383,567	107,315	1,045,850	1,971	16,088,652	381,367
Less than 18	99,348	27,631	277,080	440	4,565,204	98,681
Enrolled in public school	66,867	20,135	180,452	295	2,988,770	66,605
Public school students*	65,643	19,938	175,513	376	2,956,056	65,264
School districts	20	11	29	2	322	19
Census tracts	91	28	240	2	3,361	90
Schools*	125	54	247	4	3,339	122
School district HHI	0.23	0.19	0.16	0.01	0.93	0.24
Proportion low-income						
Individuals	0.33	0.32	0.09	0.15	0.69	0.33
Less than 18	0.41	0.41	0.11	0.12	0.80	0.42
Enrolled in public school	0.40	0.39	0.11	0.09	0.79	0.40
Public school students*	0.37	0.36	0.17	0.00	0.91	0.41
Proportion non-white						
Individuals	0.20	0.14	0.18	0.01	0.94	0.21
Less than 18	0.26	0.19	0.21	0.02	0.96	0.27
Enrolled in public school	0.22	0.16	0.19	0.00	0.96	0.22
Public school students*	0.24	0.15	0.23	0.00	0.97	0.26

*Data from CCD; the "Limited" sample drops commuting zones with too many missing values in key CCD variables.
Note: 28 commuting zones are excluded from all analyses (including this table) because they have fewer than two subdivisons.

Figure 3.1: Commuting zone summary statistics.

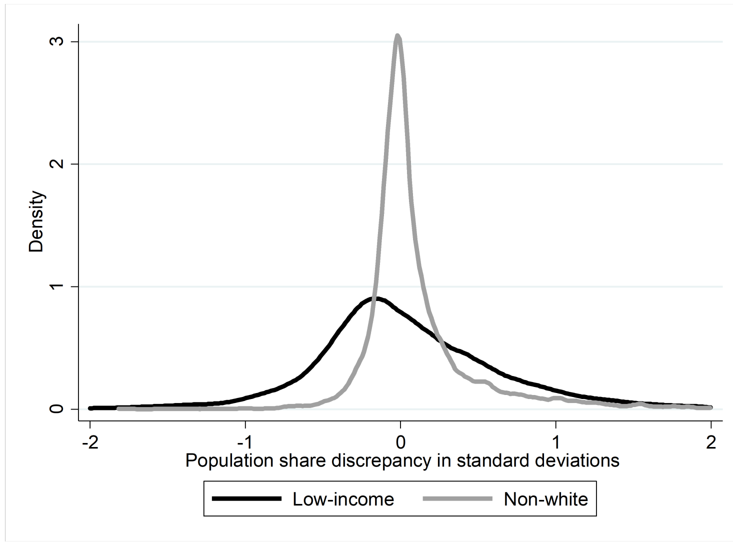


Figure 3.2: Kernel density, Discrepancy in school district population shares of CCD estimates relative to Census estimates.

included in the analysis.[3] Population characteristics exhibit right-skewed single-peaked distributions, since the population is concentrated in large commuting zones. Specifically, 50% of the total population resides in the 38 largest commuting zones, and less than 6% of the population resides in the smallest 50% of commuting zones.

CCD records are comprised of mandatory surveys of education administrative data in all public school districts and has some potential reliability issues in the 1999-2000 school year. First, data is missing for some schools. To get around this, I construct a limited sample of 588 commuting zones in which fewer than 5% of students are enrolled in schools with missing FRPL or race data, and fewer than 25% of school districts have more than 5% of students enrolled in schools with missing FRPL or race data.[4] These restrictions amount to eliminating 18% of the total population of individuals, children, and students. Fortunately, the limited sample is statistically quite similar to full sample, as shown in Figure 3.1.

The second issue with CCD counts concerns the reliability of FRPL as an income proxy. FRPL eligibility requires that household income fall below 1.85 times the federal poverty line. Yet Domina et al. (2018) document inconsistencies with income data from tax records and FRPL eligibility.

[3]28 commuting zones have fewer than two school districts and/or Census tracts, and at least two subdivisions are needed to generate segregation indices.

[4]Results are robust to different elimination thresholds, such as requiring fewer than 1% of students are enrolled in schools with missing data.

These inconsistencies arise within and across school districts. Census data offer an alternative estimate of FRPL eligible children based on income reported to the Census and household size. In the limited sample of commuting zones, on average 40% of students are FRPL eligible by Census estimates, and 41% are eligible according to CCD counts. The similarity of these figures is reassuring, but, ultimately, segregation estimates are sensitive to the distribution of the low-income population across subdivisions – not just overall rates. The discrepancies between Census and CCD estimates of FRPL eligibility are far more variable than the discrepancies in non-white population counts between the two sources. This is illustrated in Figure 3.2, which shows discrepancies in school district low-income estimates in CCD data compared to Census data, standardized relative to Census estimates.

That said, discrepancies in non-white population share estimates provide no reason to systematically prefer Census estimates over CCD counts. Unlike FRPL, there is no reason to suspect schools are inconsistent in their designations of white and minority students in their administrative records. As shown in Figure 3.1, CCD estimates of low-income and non-white shares of public school students fall between Census estimates for children enrolled in public school and all children under age 18. Potentially, Census designations for which children are enrolled in public school reflect some sort of response bias that favors white respondents, or is more accurate in less diverse commuting zones. Since I cannot favor a particular data source and have use for both tract and school subdivisions, analyses will be conducted on all relevant samples. Results will explicitly state which data source, population, sample, subgroup, and subdivision is employed for each estimate.

3.3 Measures of Segregation

Suppose that 40% of children in a commuting zone are low-income. If there were no income segregation across school districts in the commuting zone, 40% of the children enrolled in every school district would be low-income. If there were no income segregation across neighborhoods or schools, 40% of children in every neighborhood and school would also be low-income. Of course, empirically, no commuting zones have a completely even spread of low-income students. It is not unusual to encounter schools and neighborhoods where practically all students are low-income, or none are.

The measures I employ to estimate segregation within commuting zones are functions of the same basic inputs: low-income or minority population shares in each of the CZ's subdivisions, and the distribution of the CZ's population across its subdivisions. The popular Dissimilarity Index computes the share of low-income or minority students that would have to relocate in order to spread the subgroup evenly throughout the region. A second measure of income segregation – the Rank-order Variance Ratio Segregation Index – goes beyond a binary low-income indicator, and compares the full empirical income distributions of each subdivision to that of the whole region. Both measures are evenness measures, equal to zero when every subdivision has the same demographic composition, and equal to one when subgroups occupy a subset of subdivisions exclusively.

All measures are computed within commuting zones, so the total number of subdivisions will vary for each commuting zone. Three primary subdivisions are employed: Census tracts, indexed by j; school districts, indexed by k; and public schools, indexed by s. The index formulas below will use the Census tract notation, but one can equivalently swap in k or s to get the formulas for segregation across school districts and schools, respectively. Similarly, I will write "individuals," but when I turn to the empirical results, individuals may mean households with children, children (less than age 18), or public school students.

Both indices employ the same key variables:

- Q denotes the share of the commuting zones's population that belongs to the subgroup of interest (i.e., low-income or minority).
- c_j denotes the share of the commuting zone's population that resides in tract j.
- q_j denotes the share of the population of tract j that belongs to the subgroup of interest.

Q and q_j will have an additional dimension in the Rank-order Variance Ratio Segregation Index as they are computed for a range of income subgroups. Accordingly, the notation will be augmented in the description of that index.

3.3.1 Dissimilarity Index

The Dissimilarity Index is calculated for a single binary subgroup. For the Low-income Dissimilarity Index, Q and q_j represent the share of the population that is low-income in the commuting zone and within each CZ tract, respectively.[5] For the Non-white Dissimilarity Index, Q and q_j represent population shares of non-white individuals.[6]

[5] All segregation indices calculated in this paper are calculated for commuting zones. The "CZ" subscript on each variable is supressed to make the notation less cumbersome.

[6] Non-white Dissimilarity is not meaningfully different from White Dissimilarity. In fact, the data used to calculate the non-white index will rely on counts of white children, students, etc. The choice to frame it around

The Dissimilarity Index (Massey & Denton, 1989) is defined

$$D^j = \frac{1}{2Q(1-Q)} \sum_{j=1}^{J} c_j |q_j - Q| \tag{3.1}$$

Again, D^j measures the proportion of the subgroup population in a commuting zone that would have to relocate to create an even spread of the subgroup across tracts. Note that – regardless of the underlying distribution of the general population across tracts (the c_j's) – D^j equals zero when a subgroup's population share in each tract (q_j) matches its population share in the whole commuting zone (Q). An important property of D^j (and of evenness measures in general) is that it defines segregation relative to Q, so segregation can be compared across regions even when Q differs substantially.

3.3.2 Rank-order Variance Ratio Segregation Index

With sufficiently fine-grained income data, the Rank-order Variance Ratio Income Segregation Index offers a way to evaluate the evenness of the distribution of household income within a CZ across the income distribution (Reardon, 2011). This avoids relying on an arbitrary cut-off for identifying households as low-income (as required for the Low-income Dissimilarity Index). However, the complexity of this index makes its interpretation difficult.

Consider some income level $y_m \in [0, \infty)$. Q_m measures the share of households in the region that have income less than y_m; $q_{m,j}$ measures the share of households in subdivision j that have income less than y_m.

The Variance Ratio Income Segregation Index (VRSI) is defined – for a fixed y_m – as

$$R(q_m, c) = 1 - \frac{1}{Q_m(1-Q_m)} \sum_{j=1}^{J} c_j q_{m,j}(1 - q_{m,j}) \tag{3.2}$$

where $q_m = (Q_m, q_{m,1}, \ldots, q_{m,J})$ and $c = (c_1, \ldots, c_J)$. Note that the numerator and denominator utilize the binomial variance formula, so $R(q_m, c)$ is a measure of how the population-weighted average variance of income above and below y_m across tracts compares to the variance of income above and below y_m for the commuting zone as a whole. Variance is maximized at $q_{m,j} = .5$, where a community has as many households with income above y_m as it has households with income

non-white students was to maintain interpretive consistency with the low-income index – low-income and non-white children/students constitute less than half of the population in the vast majority of commuting zones. So both dissimilarity indices can be interpreted as the smallest share of students in a region that would need to be relocated to establish an even population spread.

below y_m. The income segregation index is high when a lot of subdivisions are low-variance or the population is concentrated in low-variance subdivisions. Income segregation is low when the population is concentrated in high-variance subdivisions or subdivisions generally are high-variance.

The description of $R(q_m, c)$ above assumes a fixed value of y_m. This is a useful formulation for empirical applications since generally the income distribution is observed at only a handful of income levels. Household income from the 2000 Census is reported in 16 bins, so $m \in 1, \ldots, 15$ with $y_1 = \$10,000$ and $y_{15} = \$200,000$. Of course, calculating $R(q_m, c)$ at each threshold yields 15 segregation measures for each commuting zone. To arrive at a single value for each commuting zone, the Rank-order VRSI integrates over the range of measures (Reardon 2011).

Rank-order VRSI is derived by using a polynomial approximation to evaluate $R(q_m; c)$ over its theoretical support ($q \in (0, 1)$) and integrating:

$$\Lambda^j = \frac{3}{2} \int_0^1 q(1 - q)R(q, \bar{c})dq. \tag{3.3}$$

When $R(q_m, c)$ is evaluated at a single income cut-off, it is sensitive to the underlying income composition of the region, unlike the Dissimilarity Index. If two communities have similar distributions of low-income households across subdivisions, but one has more low-income households overall (specifically, the low-income cutoff is closer to the median income), the poorer region will register a higher $R(q_m, c)$. However, in the Rank-order VRSI formulation (Λ^j), where the full range of income cut-offs are evaluated, this sensitivity washes out, because the observed range of income cut-offs offer good coverage of the full income distribution in all commuting zones and subdivisions.

3.3.3 Segregation estimates

To arrive at a single summary figure for each segregation index over each geographical subdivision in Figure 3.3, I calculate mean segregation weighted by commuting zone population. Figure 3.3 also offers a useful summary of the data and geographic coverage available for each measure. The first two columns reflect segregation estimates that include all 710 commuting zones, while the right three columns are restricted to the 588 commuting zones for which CCD data are reliable in the 1999-2000 school year. Segregation measures based on administrative counts of public school student are calculated only for the limited CZ sample. Within a row in Figure 3.3, one can see how the level of segregation depends on the subdivision being used. For example, the Rank-order VRSI of households with children across tracts (.162) is more than double its value across school districts (.067). Within a column, one can compare all the segregation indices calculated at that

	All commuting zones (n=710)		Limited (n=588)		
	SD	Tract	SD	Tract	School
Rank-order VRSI					
Households	0.040	0.099	0.040	0.100	-
Households with children	0.067	0.162	0.066	0.163	-
Enrolled in public school	0.074	-	0.073	-	-
Low-income Dissimilarity					
All individuals	-	0.311	-	0.312	-
Children less than 18	-	0.370	-	0.372	-
Enrolled in public school	0.258	-	0.255	-	-
Public school students*	-	-	0.337	-	0.454
Non-white Dissimilarity					
All individuals	-	0.471	-	0.474	-
Children less than 18	-	0.506	-	0.509	-
Enrolled in public school	0.386	-	0.378	-	-
Public school students*	-	-	0.451	-	0.524

*Data from CCD; the "Limited" sample drops commuting zones with too many missing values in key CCD variables.

Note: all means weighted by commuting zone population.

Figure 3.3: Mean commuting zone segregation indices.

subdivision level. For example, Non-white Dissimilarity for children across tracts (.506) exceeds Low-income Dissimilarity for children across tracts (.370).

Evaluating the extent of segregation is subjective. Massey & Denton (1989) selected racial dissimilarity index values greater than .6 as "high," arguing that such levels imply a qualitatively different experience of segregation than moderate and low levels. By that standard, 12% of commuting zones (in the limited sample) exhibit high levels of racial segregation across schools, with the population-weighted mean index at a moderate – but substantial – .52. This implies that just over half of minority students, on average, would have to switch schools to create an even and proportional spread of minority students across schools. Residential racial segregation across tracts for all children (not just public school students) comes in nearly as high, with a weighted mean of .51.

Using these thresholds to evaluate income segregation, fewer than 1% of commuting zones are classified as having high income segregation across schools, with a weighted mean index of .45 for public school students. This contrasts to the potentially more reliable estimate of children's Low-income Dissimilarity Index across tracts, which averages .37. However, the .6 threshold has not been validated for income segregation, and since income is inherently more variable and potentially less visible than race, the appropriate threshold for designating regions as highly income segregated may be lower. Notably, Chetty et al. (2014) use similar segregation measures over the same geographies and find the negative correlation between income segregation and income mobility is higher than that between racial segregation and income mobility. Their analysis also yielded racial segregation indices that were higher and more variable than the income segregation measures.[7]

One pattern that holds in nearly all commuting zones is that children experience higher levels of residential segregation compared to households generally (consistent with the findings of Owens 2016). This is particularly true for income segregation, and, as the Rank-order VRSI estimates suggest, the disparity may be even greater for public school-attending children.

3.4 Methodology

School districts vary in size regionally. Their borders trace counties throughout much of the South and West, but towns and cities in most of the Northeast and Midwest. This stands in contrast to Census tracts, which aim to include about 4,000 individuals regardless of region. Consider the

[7]See Online Table 8 from Chetty et al. (2014) available on opportunityinsights.org.

commuting zones containing Boston, MA and Miami, FL. In the 1999-2000 school year, the Boston CZ enrolled approximately 750,000 public school students in 264 school districts, while the Miami CZ had approximately 610,000 students in just three school districts. But the number of districts does not tell the whole story. Even where school districts are decentralized, center city school districts educate a large share of students. Roughly one-in-twelve students in the Boston CZ were enrolled the center city school district in 1999-2000.

If households are sorting in reaction to school district-level demographics and school quality, having less concentrated school districts may lead to higher segregation. Parents may be able to more clearly discern the community and school district attributes they prefer when there are many small school districts to compare. Or to frame the hypothesis in terms of highly concentrated school districts: children usually attend school close to where they live. When school districts are large, children residing in different areas of the district likely attend different schools. Parents may care more about which schools are nearby or which catchment area they live in as opposed to which school district they live in.

To test whether school district segregation is decreasing in school district concentration, I run the following regression:

$$SEG^k_{CZ} = \beta_0 + \beta_1 \log(N_{CZ}) + \beta_2 HHI^k_{CZ} + \epsilon_{CZ} \tag{3.4}$$

The Herfindahl-Hirschman index measures school district concentration – or fragmentation as Owens (2016) operationalizes it (recall, school districts are indexed by k):

$$HHI^k = \sum_{k=1}^{K} c_k^2 \tag{3.5}$$

HHI takes on low values when there are many similarly-sized school districts, and higher values when students are concentrated into a couple large districts. For example, in the 1999-2000 school year, the Boston CZ had an HHI of 0.01, while the Miami CZ had an HHI of 0.50 (see Figure 3.1 for a summary of HHI nationatlly). Including commuting zone population (N_{CZ}) in the above regression ensures that the effect of school district concentration is separate from the effect of large commuting zones, which also may facilitate higher levels of segregation by increasing the number of subdivisions, generally. If β_2 is negative and significant, segregation across school districts decreases when school districts are more concentrated.

Yet school districts do not necessarily reflect the communities children experience at home and in school. Especially when they operate over large geographic regions. If school district boundaries are the primary drivers of child segregation, then segregation measured across schools and neighborhoods should react to school district concentration in the same manner as school district segregation. To see whether this is the case, I run the same regression above, but swap in school and tract-level segregation indices in place of school district ones. In this second set of regressions, the sign and significance of β_2 tests how school and neighborhood sorting corresponds to *school district* concentration.

The regressions above offer interesting descriptive results, but are not necessarily practical for policy, since these geographical attributes – CZ size and school district concentration – are not easily augmented. To understand how sorting by income and race flow into the local communities children experience, the subsequent analysis decomposes national variation in school and neighborhood subgroup population shares across nested geographies. As in Rivkin (1994), this analysis measures the extent to which variation in subgroup populations occurs within or across school districts, but takes an approach that mirrors the geographic decomposition of Chetty et al. (2018).

Recall from the segregation formulas that segregation is a function of the relative presence of subgroups in different subdivisions (i.e., the variation in q_j's, q_k's, etc.). By decomposing the variation in subgroup population shares across nested geographies, I can identify the geographies most salient to income and racial segregation. While the segregation indices are calculated within commuting zones in order to make comparisons across them, the following analysis of the population shares of tracts throughout the country seeks to characterize the geographies over which children are sorted. This starts by estimating the share of population variation across tracts that is attributable to sorting across commuting zones. Consider the following weighted least squares regression:

$$q_{j,CZ} = \beta_0 + \gamma_{CZ} + \epsilon_j \tag{3.6}$$

where weights are the population of each tract and γ_{CZ} are commuting zone fixed effects. The resulting adjusted r-squared characterizes the variation in subgroup shares across tracts that are explained by sorting across commuting zones.

Subsequently, I swap in county-level fixed effects, which subsume commuting zone fixed effects. The resulting r-squared characterizes the variation in subgroup shares across tracts that are explained by sorting across both counties and commuting zones. The variation left over pertains to sorting across tracts within a county.

I repeat the same analysis across schools. In addition to commuting zone and counties, schools are nested into school districts, which are largely subsets of counties. Thus, for the school analysis,

Index	Rank-order VRSI	Rank-order VRSI	Low-income D
Subdivision	SD	Tract	Tract
Log(CZ pop.)	0.007***	0.032***	0.051***
	(0.001)	(0.001)	(0.002)
SD HHI	-0.044***	0.066***	0.104***
	(0.005)	(0.008)	(0.016)
N	710	710	710
R-sq	0.32	0.69	0.57

Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Note: Rank-order VRSI indices based on households with children. Low-income Dissimilarity Index calculated for children under age 18.

Figure 3.4: Linear regression, Income segregation of children less than 18 years old.

I can additionally compare the variation in subgroup population shares explained across school districts, to the variation left over across schools within school districts. It is this latter estimate – the variation in sorting across schools within districts – that bounds the capacity of school districts to augment segregation across schools.

3.5 Findings

For households with children, school district income segregation is decreasing in school district HHI, as show in the first column of Figure 3.4. However, when income segregation is measured across Census tracts – not school districts – income segregation is increasing in HHI. That is, households with children are *more* segregated across neighborhoods when school districts are concentrated. This pattern holds whether income segregation is measured by Rank-order VRSI (column 2) or Low-income Dissimilarity (column 3). This does not necessarily mean that schools are not relevant to the local sorting of households with children. Potentially, school location and catchment areas play an important role for residential income segregation in large school districts. District boundaries simply are not relevant enough to capture the relationship.

Turning the focus from all children to only those enrolled in public school, I perform the same analysis for Low-income and Non-white Dissimilarity Indices for school districts and schools (see

	Low-income Dissimilarity Index				Non-white Dissimilarity Index			
Source	Census	Census	CCD	CCD	Census	Census	CCD	CCD
Subdivision	SD	SD	SD	School	SD	SD	SD	School
Sample	Full	Limited	Limited	Limited	Full	Limited	Limited	Limited
Log(CZ pop.)	0.016***	0.016***	0.018***	0.044***	0.005	0.005	0.009*	0.023***
	(0.003)	(0.003)	(0.004)	(0.002)	(0.004)	(0.004)	(0.005)	(0.004)
SD HHI	-0.173***	-0.190***	-0.250***	0.063***	-0.479***	-0.486***	-0.487***	-0.253***
	(0.019)	(0.019)	(0.026)	(0.022)	(0.034)	(0.036)	(0.039)	(0.037)
N	710	588	588	588	708	586	588	588
R-sq	0.25	0.28	0.21	0.41	0.25	0.25	0.24	0.17

Robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Note: Indices calculated for children in public school.

Figure 3.5: Linear regression, Income and racial segregation of public school students.

Figure 3.5). The findings for the Low-income Dissimilarity Index mirror those for all children: school district income segregation is decreasing in school district concentration, but *school* income segregation is increasing in school district concentration. Income segregation among schools persists when school districts are large.

For racial segregation, the findings differ. School racial segregation decreases with school district concentration. This could reflect concerted efforts within school districts to racially desegregate schools. It could also reflect persistent racial segregation across school district boundaries, some of which has been attributed to white flight following desegregation efforts (Reber 2005).[8]

The differing geography of racial and income segregation suggests equating the two carries risks. Suppose school districts are asked to reduce racial and income segregation across their schools. If segregation is driven by behavior across – and not within – school district boundaries, then school districts will have limited ability to affect change in school segregation.

To investigate this issue further, I decompose the variation in subgroup population shares across schools (i.e., the q_s's in the segregation indices) that is attributable to larger geographical units

[8] School desegregation efforts were not the start of white flight. Shertzer & Walsh (2016) and Boustan (2010) document white flight prior to Brown v. Board of Education and subsequent court desegregation orders. However, school desegregation efforts were mostly contained within districts and, thus, may have affirmed school district boundaries as the relevant geographies across which to flee.

	Low-income population shares			Non-white population shares		
Source	Census	Census	CCD	Census	Census	CCD
Subdivision	Tracts	Tracts	Schools	Tracts	Tracts	Schools
Sample	Full	Limited	Limited	Full	Limited	Limited
CZ fixed effects	0.174	0.183	0.185	0.371	0.371	0.413
County fixed effects	0.285	0.298	0.342	0.519	0.524	0.607
SD fixed effects			0.572			0.822
Remaining variation			*0.428*			*0.178*

Note: The Census population shares include all children under the age 18; the CCD population shares reflect only children enrolled in public school. Regression weights are population of subdivision.

Figure 3.6: Adjusted r-squared, Linear regression of population shares on geographic fixed effects.

– commuting zones, counties, and school districts. I also decompose subgroup population shares across tracts (q_j's) by geographical units, though the tract shares reflect all children, as opposed to public school-attending children. As shown in Figure 3.6, in the limited CZ sample, variation in subgroup populations across counties and commuting zones explains 30% of the variation in low-income population shares across tracts, and 34% of the variation in low-income population shares across schools. When I further subsume variation across school districts, I find that 43% of variation in the low-income population share across schools occurs within school districts.

Proceeding with the same analysis for non-white population shares, I find that 52% of variation in non-white population shares across tracts is explained by sorting at the county and commuting zone levels. This figure is 61% for the non-white population shares across schools. The geographical difference in low-income and non-white sorting jumps out once school district fixed effects are added to the school-level analysis: only 18% of the variation in school non-white population shares occurs within school districts.

Since CCD data is collected annually, I can calculate these figures each year. Figure 3.7 shows that the proportion of subgroup population shares that occur across school districts, counties, and commuting zones is quite stable. After the 1999-2000 school year, low-income variation accounted for by school districts and larger geographies hovers between 57 and 62%. The figure for non-white variation is even more stable, with 81-83% of variation in non-white school populations occurring across school district boundaries. On a national scale, schools within districts exhibit substantially

	Low-income	Non-white	No. of CZs
1998	0.516	0.795	467
1999	0.572	0.822	575
2000	0.570	0.823	564
2001	0.573	0.817	547
2002	0.580	0.819	558
2003	0.571	0.812	529
2004	0.594	0.809	538
2005	0.613	0.828	574
2006	0.581	0.817	563
2007	0.603	0.821	549
2008	0.619	0.826	561
2009	0.606	0.817	567
2010	0.609	0.820	585
2011	0.614	0.811	562
2012	0.597	0.817	563

Note: Analysis restricted to limited CZ sample, with additional exclusions made each year on the basis of missing data.

Figure 3.7: Adjusted r-squared, Linear regression of annual school enrollment shares on school district fixed effects.

more income variation than they do racial variation.

3.6 Discussion

Segregation offers a potential mechanism through which racial and income disparities in outcomes foment within regions, as subgroups differentially access quality neighborhoods and schools. Not only do segregated regions exhibit lower income mobility, but segregation has also been linked to entrenched educational achievement gaps (Chetty et al. 2014; Vigdor & Ludwig 2007; Card & Rothstein 2007). Widespread school racial desegregation efforts of the 1970s and 1980s are causally linked to narrowing racial gaps in test scores, high school completion, and labor market outcomes (Guryan 2004; Ashenfelter et al. 2005).

Understanding the differing geography of racial and income segregation is critical for understanding the potential causal relationship between racial segregation, income segregation, and inequality. Moreover, the geographical variations in racial and income sorting documented here may explain why convergence in one disparity might not coincide with convergence in the other. Variation in resources, educational quality, and peer groups across school districts and larger regions may drive the racial achievement gap, while variation in these factors across schools within districts may contribute to the income achievement gap. From a policy perspective, variation in subgroup populations within school districts can be altered via school assignment policies, but the legacy of *Milliken v. Bradley* prevents policy measures that seek desegregation across school district lines. Individual school districts have greater capacity to reduce income segregation within their boundaries than racial segregation.

References

Aizer, Anna. 2008. "Peer Effects and Human Capital Accumulation: The Externalities of ADD." Cambridge, MA.

Aizer, Anna, Janet Currie, Peter Simon, and Patrick Vivier. 2018. "Do Low Levels of Blood Lead Reduce Children's Future Test Scores?" American Economic Journal: Applied Economics 10 (1): 307–41.

Aizer, Anna, Laura Stroud, and Stephen Buka. 2009. "Maternal Stress and Child Well-Being: Evidence from Siblings."

Allison, Paul D., and Richard P. Waterman. 2002. "7. Fixed-Effects Negative Binomial Regression Models." Sociological Methodology 32 (1): 247–65.

Ashenfelter, Orley, William J. Collins, and Albert Yoon. 2005. "Evaluating the Role of Brown vs. Board of Education in School Equalization, Desegregation, and the Income of African Americans." NBER Working Paper.

Avvisati, F., M. Gurgand, N. Guyon, and E. Maurin. 2014. "Getting Parents Involved: A Field Experiment in Deprived Schools." The Review of Economic Studies 81 (1): 57–83.

Balu, Rekha. 2019. "Intervention Design Choices and Evaluation Lessons from Multisite Field Trials on Reducing Absenteeism." In Absent From School, 199–212. Cambridge, MA: Harvard Education Press.

Bifulco, Robert, and Helen F. Ladd. 2007. "School Choice, Racial Segregation, and Test-Score Gaps: Evidence from North Carolina's Charter School Program*." Journal of Policy Analysis and Management 26 (1): 31–56.

Boustan, Leah Platt. n.d. "Was Postwar Suburbanization 'White Flight'? Evidence from the Black Migration." Quarterly Journal of Economics, 27.

Cabus, Sofie J., and Kristof De Witte. 2015. "The Effectiveness of Active School Attendance Interventions to Tackle Dropout in Secondary Schools: A Dutch Pilot Case." Empirical Economics 49 (1): 65–80.

Card, David, and Jesse Rothstein. 2007. "Racial Segregation and the Black–White Test Score Gap." Journal of Public Economics 91 (11–12): 2158–84.

Carrell, Scott E., and Mark Hoekstra. 2014. "Are School Counselors an Effective Education Input?" Economics Letters 125 (1): 66–69.

Carrell, Scott E, and Mark L Hoekstra. 2010. "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids." American Economic Journal: Applied Economics, January, 19.

Case, Anne, Angela Fertig, and Christina Paxson. 2005. "The Lasting Impact of Childhood Health and Circumstance." Journal of Health Economics 24 (2): 365–89.

Chetty, Raj, John Friedman, Nathaniel Hendren, Maggie Jones, and Sonya Porter. 2018. "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility." w25147. Cambridge, MA: National Bureau of Economic Research.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States *." The Quarterly Journal of Economics 129 (4): 1553–1623.

Clotfelter, Charles T. 1999. "Public School Segregation in Metropolitan Areas." Land Economics 75 (4): 487–504.

Cowen, Joshua. 2010. "Who Chooses, Who Refuses? Learning More from Students Who Decline Private School Vouchers." American Journal of Education 117 (1): 1–24.

Cowen, Joshua, and Marcus A. Winters. 2013. "Choosing Charters: Who Leaves Public School as an Alternative Sector Expands?" Journal of Education Finance 38 (3): 210–29.

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." AEA Papers and Proceedings 97 (2): 17.

Currie, Janet, and Cathy Spatz Widom. 2010. "Long-Term Consequences of Child Abuse and Neglect on Adult Economic Well-Being." Child Maltreatment 15 (2): 111–20.

Currie, Janet, and Erdal Tekin. 2012. "Understanding the Cycle: Childhood Maltreatment and Future Crime." The Journal of Human Resources, 42.

Cutler, David M., and Edward L. Glaeser. 1997. "Are Ghettos Good or Bad?" The Quarterly Journal of Economics 112 (3): 827–872.

Dahl, Gordon B, and Lance Lochner. 2012. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." American Economic Review 102 (5): 1927–56.

Domina, Thurston, Nikolas Pharris-Ciurej, Andrew M. Penner, Emily K. Penner, Quentin Brummet, Sonya R. Porter, and Tanya Sanabria. 2018. "Is Free and Reduced-Price Lunch a Valid Measure of Educational Disadvantage?" Educational Researcher.

Doyle, Joseph J., and Anna Aizer. 2018. "Economics of Child Protection: Maltreatment, Foster Care, and Intimate Partner Violence." Annual Review of Economics 10 (1): 87–108.

Engberg, John, Dennis Epple, Jason Imbrogno, Holger Sieg, and Ron Zimmer. 2014. "Evaluating Education Programs That Have Lotteried Admission and Selective Attrition." Journal of Labor Economics 32 (1): 27–63.

Epple, Dennis, and Glenn J. Platt. 1998. "Equilibrium and Local Redistribution in an Urban Economy When Households Differ in Both Preferences and Incomes." Journal of Urban Economics 43: 23–51.

Epple, Dennis, and Richard E. Romano. 1998. "Competition between Private and Public Schools, Vouchers, and Peer-Group Effects." American Economic Review 88 (1): 33–62.

Fantuzzo, John W., Whitney A. LeBoeuf, and Heather L. Rouse. 2014. "An Investigation of the Relations Between School Concentrations of Student Risk Factors and Student Educational Well-Being." Educational Researcher 43 (1): 25–36.

Fantuzzo, John W., Staci M. Perlman, and Erica K. Dobbins. 2011. "Types and Timing of Child Maltreatment and Early School Success: A Population-Based Investigation." Children and Youth Services Review 33 (8): 1404–11.

Garcia, David R. 2008. “The Impact of School Choice on Racial Segregation in Charter Schools.” Educational Policy 22 (6): 805–29.

Gershenson, Seth, Alison Jacknowitz, and Andrew Brannegan. 2017. “Are Student Absences Worth the Worry in U.S. Primary Schools?” Education Finance and Policy 12 (2): 137–65.

Goodman, Joshua. 2014. “Flaking Out: Student Absences and Snow Days as Disruptions of Instructional Time.” w20221. Cambridge, MA: National Bureau of Economic Research.

Gottfried, Michael A. 2014. “Can Neighbor Attributes Predict School Absences?” Urban Education 49 (2): 216–50.

Guryan, Jonathan. 2004. “Desegregation and Black Dropout Rates.” American Economic Review 94 (4): 919–43.

Kenney, Genevieve M, Victoria Lynch, Michael Huntress, Jennifer M Haley, and Nathaniel Anderson. 2012. “Medicaid/CHIP Participation Among Children and Parents.” Timely Analysis of Immediate Health Policy Issues, December, 15.

Krueger, Alan B., and Diane M. Whitmore. 2001. “The Effect of Attending a Small Class in the Early Grades on College-test Taking and Middle School Test Results: Evidence from Project Star.” The Economic Journal 111 (468): 1–28.

Lazear, Edward P. 2001. “Educational Production.” The Quarterly Journal of Economics.

Maloney, Tim, Nan Jiang, Emily Putnam-Hornstein, Erin Dalton, and Rhema Vaithianathan. 2017. “Black–White Differences in Child Maltreatment Reports and Foster Care Placements: A Statistical Decomposition Using Linked Administrative Data.” Maternal and Child Health Journal 21 (3): 414–20.

Manski, Charles F. 1993. “Identification of Endogenous Social Effects: The Reflection Problem.” The Review of Economic Studies 60 (3): 531.

Massey, Douglas S., and Nancy A. Denton. 1989. “Hypersegregation in U.S. Metropolitan Areas: Black and Hispanic Segregation along Five Dimensions.” Demography 26 (3): 373.

Michelmore, Katherine, and Susan Dynarski. 2017. "The Gap Within the Gap: Using Longitudinal Data to Understand Income Differences in Educational Outcomes." AERA Open 3 (1).

Moffitt, Robert A. 2001. "Policy Interventions, Low-Level Equilibria and Social Interactions." In Social Dynamics. MIT Press.

Nechyba, Thomas. 2003. "School Finance, Spatial Income Segregation, and the Nature of Communities." Journal of Urban Economics 54 (July): 61–88.

Owens, Ann. 2016. "Inequality in Children's Contexts: Income Segregation of Households with and without Children." American Sociological Review 81 (3): 549–74.

———. 2018. "Income Segregation between School Districts and Inequality in Students' Achievement." Sociology of Education 91 (1): 1–27.

Pieterse, Duncan. 2015. "Childhood Maltreatment and Educational Outcomes: Evidence from South Africa." Health Economics 24 (7): 876–94.

Reardon, Sean F. 2011. "Measures of Income Segregation."

Reardon, Sean F., Elena Tej Grewal, Demetra Kalogrides, and Erica Greenberg. 2012. "Brown Fades: The End of Court-Ordered School Desegregation and the Resegregation of American Public Schools: Brown Fades." Journal of Policy Analysis and Management 31 (4): 876–904.

Reardon, Sean F., Demetra Kalogrides, and Kenneth Shores. 2017. "The Geography of Racial/Ethnic Test Score Gaps." SSRN Electronic Journal.

Reback, Randall. 2010. "Schools' Mental Health Services and Young Children's Emotions, Behavior, and Learning: Schools' Mental Health Services and Young Children's Emotions." Journal of Policy Analysis and Management 29 (4): 698–725.

Reber, Sarah J. 2005. "Court-Ordered Desegregation: Successes and Failures Integrating American Schools since Brown versus Board of Education." Journal of Human Resources XL (3): 559–90.

Rivkin, Steven G. 1994. "Residential Segregation and School Integration." Sociology of Education 67 (4): 279.

Rocque, Michael, Wesley G. Jennings, Alex R. Piquero, Turgut Ozkan, and David P. Farrington. 2017. "The Importance of School Attendance: Findings From the Cambridge Study in Delinquent Development on the Life-Course Effects of Truancy." Crime & Delinquency 63 (5): 592–612.

Romero, Mariajosé, and Young-Sun Lee. 2007. "A National Portrait of Chronic Absenteeism in the Early Grades." National Center for Children in Poverty.

Ryan, Joseph P., Brian A. Jacob, Max Gross, Brian E. Perron, Andrew Moore, and Sharlyn Ferguson. 2018. "Early Exposure to Child Maltreatment and Academic Outcomes." Child Maltreatment 23 (4): 365–75.

Shertzer, Allison, and Randall P. Walsh. 2016. "Racial Sorting and the Emergence of Segregation in American Cities." NBER Working Paper.

Slade, Eric P., and Lawrence S. Wissow. 2007. "The Influence of Childhood Maltreatment on Adolescents' Academic Performance." Economics of Education Review 26 (5): 604–14.

Smythe-Leistico, Kenneth, and Lindsay C. Page. 2018. "Connect-Text: Leveraging Text-Message Communication to Mitigate Chronic Absenteeism and Improve Parental Engagement in the Earliest Years of Schooling." Journal of Education for Students Placed at Risk (JESPAR) 23 (1–2): 139–52.

Sutphen, Richard D., Janet P. Ford, and Chris Flaherty. 2010. "Truancy Interventions: A Review of the Research Literature." Research on Social Work Practice 20 (2): 161–71.

Vigdor, Jacob, and Jens Ludwig. 2007. "Segregation and the Black-White Test Score Gap." w12988. Cambridge, MA: National Bureau of Economic Research.