ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 19-86

The Effects of Special Education on the Academic Performance of Students with Learning Disabilities

Amy Ellen Schwartz
Syracuse University

Bryant Gregory Hopkins
New York University

Leanna Stiefel
New York University

In the forty plus years since passage of the *Individuals with Disabilities Education Act* (IDEA), special education has grown in the number of students and amount spent on services. Despite this growth, the academic performance of students with disabilities (SWDs) remains troubling low compared to general education students (GENs). To some extent, these differences reflect persistent underlying disabilities, but they may also reflect ineffective special education services. Does special education improve academic outcomes for students with disabilities? There is surprisingly little evidence to guide policy and answer this question. This paper provides an answer for the largest disability group, students with learning disabilities (LDs), using rich data from New York City public schools. Because the majority of LDs are classified after school entry, we observe outcomes both before and after classification, allowing us to gauge impact using within-student pre/post comparisons and, ultimately, student fixed effects in regression models exploring impacts. We find that academic outcomes improve for LDs following classification into special education, and impacts are largest for those entering special education in earlier grades. Results are robust to alternative specifications and falsification tests bolster confidence in a causal interpretation. Differences in impacts by gender and race/ethnicity, grade of classification, and settings shed light on possible mechanisms.

VERSION: June 2019

Suggested citation: Schwartz, Amy Ellen, Bryant Gregory Hopkins, Leanna Stiefel. (2019). The Effects of Special Education on the Academic Performance of Students with Learning Disabilities. (EdWorkingPaper: 19-86). Retrieved from Annenberg Institute at Brown University: http://www.edworkingpapers.com/ai19-86

The Effects of Special Education on the Academic Performance of Students with Learning Disabilities

June 16, 2019

Amy Ellen Schwartz
Maxwell School, Syracuse University
Professor of Economics and Public Administration and International Affairs, Daniel Patrick Moynihan Chair in Public Affairs
amyschwartz@maxwell.syr.edu

Bryant Gregory Hopkins
Wagner School, New York University
Ph.D. student and IES PIRT fellow
bryant.hopkins@nyu.edu

Leanna Stiefel
Wagner School, New York University
Professor of Economics
leanna.stiefel@nyu.edu

We thank the New York City Department of Education for access to data and technical support, and participants at the 2018 Association for Education Finance and Policy conference, the 2018 Association for Policy Analysis and Management conference, the 2019 American Economic Association Conference, the NYU-PRIISM summer seminar series, and the Syracuse University Education and Social Policy workshop for helpful comments. Generous support from the Spencer Foundation (grant 201500141) and Institute for Education Sciences sponsored PIRT fellowship program are gratefully acknowledged. Authors alone are responsible for all results.

Abstract

In the forty plus years since passage of the *Individuals with Disabilities Education Act* (IDEA), special education has grown in the number of students and amount spent on services. Despite this growth, the academic performance of students with disabilities (SWDs) remains troubling low compared to general education students (GENs). To some extent, these differences reflect persistent underlying disabilities, but they may also reflect ineffective special education services. Does special education improve academic outcomes for students with disabilities? There is surprisingly little evidence to guide policy and answer this question. This paper provides an answer for the largest disability group, students with learning disabilities (LDs), using rich data from New York City public schools. Because the majority of LDs are classified after school entry, we observe outcomes both before and after classification, allowing us to gauge impact using within-student pre/post comparisons and, ultimately, student fixed effects in regression models exploring impacts. We find that academic outcomes improve for LDs following classification into special education, and impacts are largest for those entering special education in earlier grades. Results are robust to alternative specifications and falsification tests bolster confidence in a causal interpretation. Differences in impacts by gender and race/ethnicity, grade of classification, and settings shed light on possible mechanisms.

I. Introduction

In the forty plus years since passage of the *Individuals with Disabilities Education Act* (IDEA), special education has grown, both in the number of students enrolled and amount spent on services, aiming to provide a "free and appropriate public education" to children with disabilities.[1,2] Despite this growth, the academic performance of students with disabilities (SWDs) is troublingly low. According to the 2017 National Assessment of Educational Progress (NAEP), only 16 (12) percent of SWDs were proficient in math (reading) as of grade 4, significantly lagging the 44 (40) percent of general education students (GENs).[3] At the same time, SWD high school graduation rates lag GENs' by 14 percentage points (see Figure 1).[4] To some extent, these differences reflect underlying disabilities, which can include cognitive impairments that limit academic performance regardless of supports or services. That said, the disparities may also reflect ineffective special education services. Does special education work to improve academic outcomes for students with disabilities? This paper provides an answer to this question for the largest disability group - students with learning disabilities (LDs). Specifically, we use longitudinal, student-level data on more than 44,000 LDs over seven years to derive credibly causal estimates of the effect of special education services on academic performance.

---

[1] Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004). The legislation was originally titled the *Education of Handicapped Children Act.* The name was changed in 1990.

[2] In 1970, there were 2.7 million students with disabilities, representing 5.9% of public school enrollment (Snyder, 1993). By 2015, that number had risen to 6.6 million, representing 13% of enrollment (USDOE, 2018). Per-pupil spending on special education increased from $1,257 to $12,474 between 1968-69 and 1999-2000 (Parrish *et al.*, 2003).

[3] In 8th grade, only 9 (10) percent of SWDs were proficient in math (reading), compared to 38 (40) percent for GENs (National Center for Education Statistics, 2018).

[4] Additionally, in 2014-15, 18.9 percent of SWDs and 12.9 percent of GENs missed at least 10 percent of school instruction, SWDs lag behind GENs in employment, income, and life satisfaction, and lead on negative indicators including suspensions, expulsions, and delinquency (Blackorby & Wagner, 1996; The Condition of Education, 2018; Glander, 2016; Wagner et al., 2006; Phelps & Hanley-Maxwell, 1997).

There is surprisingly little evidence to guide special education policy and answer the question of whether services work. Although there is a large and growing literature estimating the effects of school-based policies and reforms on academic and non-academic outcomes among GENs, the estimation of similar effects for SWDs is quite limited. This dearth of research reflects heterogeneity among SWDs, endogeneity of service receipt, and data scarcity. SWDs are heterogeneous, grouped by the federal government into 13 classifications that differ in observed and unobserved ways from each other and from GENs. These differences are related to both receipt of services and outcomes, making it difficult to obtain unbiased impact estimates; moreover, solving the conceptual difficulties requires significant amounts of data that are hard to assemble. The administrative data sets used in studies of GENs typically lack SWD classifications or services for the sufficiently large sample of students or schools required for causal research designs. Our uniquely rich data from New York City (NYC) provide both the requisite variables and a large sample of SWDs.

Among all SWDs, we focus on LDs in particular for two reasons. First, LDs are the largest group of SWDs - representing 35 percent of SWDs nationally and 40 percent in NYC in 2015.[5] Second, since the majority of LDs are classified *after* school entry (typically grades 3 through 8),[6] we observe outcomes both before and after classification allowing us to gauge impact using within-student pre/post comparisons and, ultimately, student fixed effects in regression models exploring impacts. This differential timing of classification also allows us to explore heterogeneity in the treatment effect by grade of classification.

To preview the results, we find that academic outcomes improve for LDs following classification into special education – math and ELA scores increase by .117 and .102 standard

[5] USDOE, 2018

[6] 71% of NYC LDs in grades K-8 are classified in grades 3 through 8.

deviations (s.d.) on average, respectively, with larger effects on mathematics performance for girls than boys (.059 s.d. higher). There is notable heterogeneity across race/ethnicity – with the largest effects for Asians and the smallest for black students – and effects are largest for those entering special education in earlier grades. Patterns observed in event study analyses are consistent with the school-based special education classification process, and results are robust to alternative samples and specifications (i.e. including LDs with a secondary disability, and limiting the sample to students continuously enrolled for three or more academic years). Finally, LDs entering special education in earlier grades are more likely to be placed, and remain, in less restrictive service settings, which may explain heterogeneity in impact estimates. Falsification tests bolster our confidence that a causal interpretation is warranted.

The remainder of this paper is organized as follows. In section II, we review the quantitative literature on the effectiveness of special education. In section III, we provide background regarding the special education classification process and the nature of learning disabilities, and in sections IV and V, we describe the data and the study's models, respectively. In section VI, we present results, and in section VII we conclude and discuss results.

II. Previous Literature on the Effect of Special Education

A variety of meta-analyses summarize specific instructional approaches that successfully increase math[7] and reading outcomes,[8] as well as components of those approaches that predict treatment outcomes, for LDs. [9] Generally, outcomes for LDs increase when interventions include

[7] Gersten et al., 2009
[8] Morgan et al., 2012; Swanson, 1999; Swanson & Sachse-Lee, 2000
[9] Swanson, 1999; Swanson, Hoskyn, & Lee 1999

some variant of direct instruction, [10] but these studies are quite specific to a single intervention.[11] To our knowledge, only three previously published papers derive credibly causal estimates of the impact of special education, *in general,* on academic performance, arguably an especially important question for public policy.

Reynolds and Wolf (1999) follow more than 1,200 low-income students from the Chicago Longitudinal Study's 1986 kindergarten cohort to estimate the impact of special education on math and reading scores. These student-level longitudinal data include annual test scores for grades K through 6, allowing the authors to estimate value-added regressions and mitigate the effects of unmeasured characteristics (e.g. personal motivation or family influence). Controlling for prior performance, family background, and school experiences and attributes, math and reading scores do not improve for SWDs.[12] Additionally, after 3rd grade, LDs perform worse than students classified with a disability other than LD. Although innovative in their estimation strategy, the study focuses only on low-income students and disability classifications are not available prior to 3rd grade.

Hanushek, Kain, and Rivkin (2002) use student-level longitudinal data on three 4th-grade cohorts of over 750,000 unique students in Texas public schools (beginning in AY 1994) to estimate the impact of special education on math test score gains. Impacts are estimated for all

[10] According to a Swanson's (1999) interpretation of the literature, direct instruction described instructional activities that included breaking down a task into steps, administering probes, administering feedback repeatedly, providing a pictorial or diagrammatic presentation, allowing for independent practice and individually paced instruction, breaking the instruction down into simpler phases, instructing in a small group, teacher modeling skill, providing set materials at a rapid pace, providing individual child instruction, teacher asking questions, and/or teacher presenting the novel materials.

[11] Hocutt (1996) provides a comprehensive summary of earlier literature, which focuses on two distinct threads in the special education research: studies comparing (1) experiences and (2) outcomes of SWDs placed in special education versus traditional classrooms. Research in the second thread indicates slightly better outcomes for LDs served in special education classrooms. In these classrooms, math and reading outcomes slowly improve, however, gains cease when placed in traditional classrooms.

[12] With the exception of 4th-grade scores in both math and reading, SWDs receiving special education services saw negative gains in achievement at each grade level.

SWDs, by disability (i.e. LD, emotional disturbance – ED, and speech impairment - SI), and separately for students who enter and remain in special education after 4th grade (*Entry*) and for students who are in special education in 4th grade and exit sometime afterwards (*Exit*).[13] Results obtained from models with student and school-grade-year fixed effects show a positive impact of special education for all SWDs (.04 s.d.). For the *Entry* sample, the estimated effect is .08, and, compared to SIs (-.02), effects for this sample are larger for LDs (.11) and EDs (.15).[14] Similar to the Reynolds and Wolf study, data availability limit generalizability of these results. The data were collected prior to *No Child Left Behind* - which required 95% of students across all demographic subgroups, including SWDs, to participate in state testing – and math gain scores are available for only 30% of LDs in the state. ELA scores are not reported in the paper, although the authors state they are similar but smaller in magnitude, and attendance is not included as an outcome in the study.

More recently, Morgan, Frisco, Farkas, and Hibel (2010) use a sample of 363 SWDs in the Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K) and a propensity score matching design to estimate the impact of services on math and reading tests, internalizing and externalizing behaviors, and learning-related tasks (i.e. remaining attentive, persisting at a given task, and being flexible/organized) at 5th grade.[15] They report a negative impact on reading scores and an insignificant impact on math scores and gain scores across both subjects. Further, compared to GENs, special education fails to reduce internalizing or

[13] The models in this study are most like ours, but our ITT model avoids post-service exit or re-entry, which may be endogenous.
[14] For the *Exit* sample, the average impact is a precisely estimated zero. There is heterogeneity across disability, with exiting LDs, SIs, and EDs seeing a .02, -.02, and .09 change in achievement, respectively.
[15] The analysis sample is limited to students with complete information regarding outcome measures and demographic characteristics. Three unique propensity score matching strategies were used to pair the 363 SWDs to a comparable group of GENs.

externalizing problem behaviors, but positively affects learning-related behaviors. The small number of observations for SWDs precludes the investigation of heterogeneity in treatment effects across disability classifications, time of classification, etc.

We use detailed data on a large sample of students with broad coverage of math, ELA, and attendance outcomes, allowing us to implement a longitudinal design and drill deep on heterogeneity to provide policy relevant answers to questions of whether special education works and for whom.

III. The Special Education Classification Process: Background and Empirical Implications

According to IDEA (2004), students are eligible for special education if they exhibit delays in thinking and learning, understanding and using language, self-help skills, physical ability, or behavior that impairs their ability to perform academically.[16] Parents, teachers, or other school personnel can refer students for special education at school entry or any grade thereafter. Referred students undergo psychoeducational and physical examinations by psychologists or other relevant professionals to determine if services are appropriate (NYCDOE, 2018a).[17] If the initial evaluation deems a student eligible for special education, parent(s), teacher(s), and school personnel develop an Individualized Education Program (IEP), which documents the student's disability, identifies supports that will be provided, service setting(s) in which supports will be administered, and academic goals for the upcoming year.[18] IEPs are re-

[16] For example, poor eyesight could impair learning but special education would not be required if corrected with eyeglasses.
[17] Additional evaluations for speech and language needs, occupational and/or physical therapy, assistive technology, and other related services are conducted as necessary.
[18] IEPs are required for each student identified with a disability under federal legislation.

evaluated annually and services are modified, continued, or discontinued as appropriate. Importantly, parents can refuse initial evaluations or services at any point.[19]

IDEA defines 13 unique disability classifications: autism, deaf-blindness, deafness, emotional disturbance, hearing-impairment, intellectual disability, learning disability, multiple handicaps, other health impairment, orthopedic impairment, speech impairment, traumatic brain injury, and visual impairment. Learning disability (LD) is by far the most common classification, comprising approximately 41% (40%) of public school special education students nationally (in NYC) in 2006.[20,21] LDs have problems with one or more of the psychological processes involved in understanding or using language, which results in an imperfect ability to speak, think, write, spell, or do math calculations.

To understand the timing of LD classification in NYC, Figure 2 illustrates the distribution of disabilities by grade, for all SWDs between 2006 and 2012. Unlike other SWDs, the majority of LDs are classified *after entry to school.* More specifically, 97% of K-8 students classified with an LD only are identified after kindergarten and nearly 71% are classified in grades 3 through 8.

Several characteristics of the LD identification and classification process are relevant to the empirical estimation of the treatment effect of special education, as explained below.

Heterogeneity by Race and Gender

Not all referrals lead to classification and not all classifications are appropriate. If all students who are potentially eligible for special education services are referred, then some of them will not, ultimately, be classified since special education services are not warranted. At the

[19] NYC Department of Education, 2018a; NYC Department of Education, 2018b
[20] 2006 is the first year of our sample period.
[21] The five largest categories are (with national percentages from 2006): learning disability (40.8%), speech impairment (21.9%), other health impairment (8.5%), emotional disturbance (7.1%), and a group that includes the nine low-incidence classifications of deafness, deaf-blindness, visual impairment, hearing impairment, intellectual disability, multiple handicap, autism, traumatic brain injury, and orthopedic impairment (21.8%).

same time, some students may be mistakenly classified and it is possible that such misclassification may be correlated with observed – or unobserved – characteristics of students, parents, or schools. As an example, disruptive students, often boys, may be referred – and classified – inappropriately due to difficulties distinguishing disabilities from other issues. Equally troubling, black students may be referred and classified inappropriately due to racism or implicit bias, leading to disproportionality in special education placements.[22] School context may also matter to the timing and appropriateness of referral and classification – the experience and expertise of a school's teachers may hinder or enhance appropriate and timely classification and placements.[23] Further, past research suggests that common cognitive testing practices used to identify LDs (e.g. comparing ability to achievement) may be culturally biased, leading to the over-identification of minorities.[24] Finally, parents may identify needs or refuse services, reflecting community norms, social capital, or preferences, which may vary by parent and student socio-demographic characteristics. For example, researchers have repeatedly found that Asian students are under-represented in special education.[25] *Thus, we may find different impacts for boys and girls and heterogeneity across race/ethnicity groups –here, Asian, black, Hispanic or white.*

Heterogeneity by Grade of Classification

Supports may be more beneficial in earlier grades when students are learning foundational skills in literacy or numeracy. Indeed, education advocates have argued that early intervention is particularly important because learning to read in the early grades is critical to

[22] Morgan et al., 2016; Cruz & Rodl, 2018; Coutinho & Oswald, 2000
[23] Hibel, Farkas, & Morgan, 2010; Boyd, Lankford, Loeb, & Wyckoff, 2005; Hanushek et al., 1999; Ingersoll, 2001;
[24] Ford (2008)
[25] Cooc, 2016.

success in later grades when students "read to learn."[26] Additionally, the *impact of services might depend on the duration of receipt* and early identification means services can be delivered for a longer time. Finally, *the grade of classification may reflect the efficacy of parent and/or school advocacy*. Identification in elementary school may indicate an engaged or effective support system, but absent this, students may have another chance to obtain services. For example, most students in the NYC public school system experience at least one school transition between grades 3 and 8. For students not enrolled in special education prior to an elementary-middle school transition, moving schools may expose students to teachers and administrators with a greater ability to identify students with disabilities.

Academic Performance and Classification

Although low performance on standardized tests does not itself indicate need for LD classification, *declining* performance may stimulate requests by school personnel or parents for evaluation for special education in general and/or LD in particular. More generally, LD students not identified in earlier grades, and therefore not receiving services or supports, are likely to fall more and more behind as inadequate visual processing or reading skills hamper their performance. Thus, we might expect *a period of declining test scores before classification in the later grades.*

Heterogeneity by Type of Service

Once classified, LD's (as all special education students) may be provided one – or more – services in one – or more – settings as needed to constitute an appropriate education. Students may be placed in a self-contained class with other students with disabilities, or a "general education" classroom with pull-out speech or language therapy, among others, detailed below.

[26] Annie E. Casey Foundation, 2010

The specific placement is likely to depend upon an assessment of the student's specific needs, as well as the availability of resources at the school. The complexities and heterogeneity of learning disabilities suggest students may differ in their needs as well as in the effectiveness of the different settings. Therefore, treatment effects are also likely *across the different special education settings/services* that may be provided to LDs.

To summarize, several considerations are important for estimating treatment effects. First, impacts may vary across gender and race/ethnicity – due to schools, teachers, discrimination, or parental decisions that induce differential selection into LD classification. Second, impacts likely differ across grade of classification, with larger effects in earlier grades if interventions are more effective in these grades, or, more generally, there are differences in the characteristics of students classified in different grades. Third, test scores may decline before classification. Fourth, the provision of services (self-contained versus general education with supports, etc.) may differ between LDs, due to unobserved differences across referral practices, availability, or effectiveness.

IV. Data, Sample, and Descriptive Statistics

We use longitudinal student-level administrative data on all 1.3 million K-8 students in 1,500 NYC public schools in 2006-2012 provided by the NYC Department of Education (NYCDOE). These include SWD status, disability classification, and primary service setting for all students with IEPs along with identifiers for race/ethnicity, gender, nativity, English

Language Learner status (ELL),[27] free or reduced-price lunch eligibility (FRPL), grade, and school.[28,29]

In our data, SWDs are identified as having an IEP on record with the NYCDOE and SWD status is binary in each academic year. Each disability classification is coded similarly and we use this information to create multiple indicators. First, *LDC* and *nonLD* take a value of one in the years that students were classified as an LD or non-LD special education student, respectively. Second, *LD,* our main independent variable, takes a value of one in the first year a student was classified as LD and all years after initial classification, even if declassified. Finally, *LD1* through *LD7* distinguish students first classified in grades 1 through 7 – that is, the interactions between grade of classification and *LD.*

We also create a set of indicators that capture the primary service a student receives in each year. In NYC, the IEP identifies a primary service setting as the service provided for the largest proportion of time during the school day and services are grouped into four distinct categories and ranked on a continuum from "least to most restrictive."[30] The least restrictive, Related Services (RS), includes support services such as counseling, physical therapy, or speech/language therapy.[31] Special Education Teacher Support Services (SETSS) encompass supplemental instruction that helps SWDs remain in the general education classroom.[32] RS and

[27] ELL is an indicator for students who are eligible for English Language Learner services: students who do not speak English as their primary language and have a limited ability to read, speak, write, or understand English.
[28] Henceforth, academic years are denoted by the year of the spring semester. For example, the academic years 2005-06 and 2011-12 are denoted 2006 and 2012, respectively.
[29] Identifiers for grade and school are used to create two more variables, *HeldBack* and *Mobility,* which identify students that repeat their current grade level or switch schools, respectively, and are used as controls in models.
[30] NYCDOE, 2016
[31] More specifically: counseling, school health services, hearing education services, occupational therapy, physical therapy, speech/language therapy, vision education services, orientation and mobility services, and "other support" services. A student's primary assigned service is designated RS if it is the only service provided by the district.
[32] SETSS are provided at a minimum of three hours a week and a maximum of 50% of the school day. The special education teacher can provide instruction directly (to a single individual or a group no larger than eight SWDs) or indirectly (in collaboration with the student's general education teacher).

SETSS are provided either in-class or as pull-out services. Integrated Co-Teaching (ICT) is a service through which GENs and SWDs are educated by both a general and special education teacher in a single classroom.[33] Self-Contained (SC) services are provided to SWDs with similar needs in SWD-majority classrooms.[34,35]

Our primary performance measures are scores on New York State (NYS) math and English Language Arts (ELA) exams administered annually to students in grades 3 through 8. We standardize these as Z-scores (mean zero and standard deviation one) for each grade-year using the citywide mean and standard deviation, to create *Z-Math* and *Z-Read*. We also use annual attendance rates to gauge academic performance between grades K and 8.[36]

We construct two analytic samples. The test score analyses use a sample of students in grades 3 through 8 and attendance analyses use a sample of students in kindergarten through 8th grade. Both include only students who attended a NYC public school for at least two years between 2006 and 2012, have test score (attendance) data for at least one year before and after

---

[33] The number of SWDs in an ICT class may not exceed 40% of the total class register, with a maximum of 12 SWDs.

[34] Students with severe disabilities are educated either in schools serving only students with IEPs or in special classrooms sited in traditional NYC public schools. These special schools and classrooms - administered by a sub-city district referred to as "District 75" (e.g. Manhattan School for Career Development, P.S. 23 at Hillside Psych Hospital, P.S. 373 K - Brooklyn Transition Center, and P.S. 721Q – Queens Occupational Training Center) are located throughout all five NYC boroughs and, thus, are not contained within a single geographic region of the city. Over the last decade, the percentage of NYC public school SWDs educated in District 75 has declined even as the total number of SWDs has increased (Authors, 2019). This shift toward inclusion means SWDs are increasingly educated in traditional schools and provided services in more collaborative classrooms containing both SWDs and GENs. Additionally, within traditional schools, there has been a large decline in the number of SWDs receiving SC, RS, and SETSS services and a significant increase in those receiving ICT services (Between 2005-06 and 2014-15, the share of SWDs receiving ICT increased from 13.1% to 46.2% [Authors *et al*., 2019]). Overall, SWD are educated less with isolating services and more with their GEN peers.

[35] Note, for our analysis, we group students receiving either RS or SETSS together in *SETSS* as less than 1% of LDs in the analytic sample receive RS.

[36] Attendance rates are calculated by dividing the total number of days a student was present at school by the number of days registered as a student within the district that academic year. This allows for an accurate calculation of attendance for students starting, or ending, the academic year in a location outside the NYC public school system.

classification, and whose only special education classification is LD.[37,38] Our math (ELA) sample includes 24,189 (23,901) unique students across 1,246 (1,245) public elementary and middle schools, for 92,902 (90,356) total student-year observations. Our attendance sample includes 44,487 unique students across 1,330 public elementary and middle schools, totaling 197,274 student-year observations.

Table 1 provides baseline (2006) descriptive statistics for all K-8 GENs, non-LDs, and LDs.[39] As shown in column 1, NYC public school GENs have high shares of black (31.0%), Hispanic (38.7%), poor (76.7%), and foreign-born students (more than 15%), roughly evenly split between boys and girls. Non-LDs and LDs differ significantly from GENs in a variety of ways. Perhaps most striking is the over-representation of non-LD boys (more than two thirds), and under-representation of foreign-born (roughly 5 percentage points). Non-LDs are also disproportionately black (34.0%), Hispanic (44.3%), poor (81.7%), and ELL (16.7%). LDs are similar to non-LDs broadly, although somewhat less disproportionately male and native-born.

Turning to outcomes, attendance among GENs exceeds 91% with non-LD and LD rates both lagging by a few percentage points (89.2% and 89.5%, respectively). On test scores, GENs perform roughly .100 sd above the district mean in math and ELA while non-LDs score between .300 to .400 sd below average in math and ELA, respectively, and LDs perform even worse – averaging .669 below the mean in math and .749 in ELA.

[37] Test coverage is high among students in the analytic sample. Within all student-year observations, missing rates for math and ELA test scores are only 2.8% and 4.6%, respectively.
[38] We also exclude student-year observations for students coded as ungraded since only 25% are tested using the NYS Math and ELA examinations. Students are coded ungraded if they are eligible to participate in the NYS Alternate Assessment and achievement is significantly below (three or more years) the grade-level coursework in math and ELA of their GEN peers (*The Data Specialist's Guide for New York City Schools,* 2008).
[39] GENs were never classified with a disability between 2006 and 2012; non-LDs were classified with a disability other than LD between 2006 and 2012. LDs were classified with one disability, LD, between 2006 and 2012.

As for specific disabilities, SIs represent roughly 60% of non-LDs, and EDs, OHs, and Others each represent less than 20%. These proportions mirror national statistics on the distribution of disability among non-LD SWDs. Finally, roughly 16% of both non-LDs and LDs attend school in an ICT setting. In contrast, non-LDs are more likely to be in a self-contained setting (45.1% and 23.0% for non-LDs and LDs, respectively) while LDs are more likely to receive SETSS or RS (60.3% and 39.3% for LDs and non-LDs, respectively).

Following the discussion about potential differences in students classified in different grades, we examine the characteristics of students and settings by grade of classification. As shown in Table 2, there are meaningful differences in the racial/ethnic composition, with the representation of newly classified students rising from 46% to 48 % among Hispanics and 32% to almost 39% among blacks between fourth and eighth grades. The representation of white students drops from roughly 12% to 8%, and from 8% to 5% for Asians. Additionally, the share of newly classified foreign-born increases, while the share of girls and ELL students both drop. Thus, early entrants are disproportionately female, white or Asian, native-born, and ELL.

Grade 3 ELA performance also differs markedly across grade of classification – those classified in grade 4 performed almost a full standard deviation below the mean, with average performance .200 or .300 higher among those classified in grade 8. At the same time, we see some evidence of a greater reliance on self-contained settings for students classified later, rather than SETSS or ICT. Below, we explore whether and how these differences correspond to differences in impacts.

V. Models

Our baseline regression model links academic outcomes and LD status, as follows:

$$Y_{igst} = \beta_0 + \delta LDC_{igst} + X'_{igst}\beta_1 + \theta_g + \tau_t + \varphi_s + \eta_i + \varepsilon_{igst} \quad (1)$$

where Y_{igst} represents an academic outcome for student i in grade g, school s, and year t; LDC_{igst} is an indicator that takes a value of 1 if student i is classified as LD in year t, 0 otherwise; X'_{igst} is a set of time-varying student demographics (i.e. *ELL*, *FRPL*, *HeldBack,* and *Mobility*); θ_g and τ_t are grade and year fixed effects, respectively, representing common grade and year specific differences; φ_s and η_i are school and student fixed effects, respectively, capturing idiosyncratic time invariant characteristics of schools (such as size, grade span, or mission) and students (such as race, gender); and ε_{igst} is an error term with the usual properties.[40]

In this student fixed effects model, the key coefficient, δ, is identified by variation in *LD* status within students and δ indicates the difference in outcomes of LD students between the years they were classified as LD and the years they were classified as GEN such that the "counterfactual" (the GEN years) potentially includes years prior to classification and years after declassification. If the educational gains from special education last beyond the year of services, as is the goal of many remedial programs, δ will underestimate the impact because the higher post-LD performance will be part of the non-LD years. Similarly, if de-classification is endogenously driven by family and school assessments of student ability to succeed with fewer supports, then δ will be downward biased.[41] To address this, we substitute an alternative measure of LD treatment that remains on after first classification, even if the student has been de-classified. Thus, we derive an intent-to-treat estimate of the impact as the difference in performance before and after LD classification.

[40] Robust standard errors are clustered at the school level in order to correct for potential correlations in the errors among students attending the same school.

[41] Students who no longer require services are considered declassified and do not have an IEP, however, if deemed necessary, can still receive additional instruction support or modification, testing accommodations, or related services. Declassification services may continue for up to one year (NYCDOE, 2016). Alternatively, if declassification is a reflection of dysfunction, instability or dissatisfaction with services, then the estimate may be upwardly biased.

To be concrete, our intent to treat specification replaces *LDC* with *LD*:

$$Y_{igst} = \beta_0 + \delta LD_{igst} + X'_{igst}\beta_1 + \theta_g + \tau_t + \varphi_s + \eta_i + \varepsilon_{igst} \quad (2)$$

where LD_{igst} takes a value of 1 if student *i* was classified as LD in *t* or any prior year, and 0 otherwise.[42] Thus, δ captures the effect of special education for LDs, including performance in all of the post-classification years in the treatment effect.

We next investigate heterogeneity in the treatment effect by gender, race/ethnicity and grade of classification. We do so by adding in *LD1* through *LD7*, the set of interaction variables between *LD* and indicators for grade of first classification.[43] The coefficients on these variables will deliver a different treatment effect for each gender, race, and grade of classification.

We then explore trends in student performance in the periods preceding and following LD classification in an event study specification. To do so, we first extend our model estimating the impact of services by grade of classification to include a set of indicator variables for each of the years preceding and following classification. These will allow us to examine the patterns in performance prior to and after classification, shedding light on potential factors that may have led to classification and informing our interpretation of the estimated impacts. We then include an additional set of indicators to separate post-classification patterns for students that remain in special education (LDC) and those that were declassified (GEN).

Finally, we explore the relationship between LD and placement in different services/settings using event study models. Specifically, we estimate a set of models linking service indicators (ICT_{igst}, $SETSS_{igst}$, GEN_{igst}, SC_{igst}) and a set of indicators capturing the time pre and post classification.

[42] That is, *LD* identifies students who are or have been classified as LD in any previous year, identifying them as LD in all years after initial classification, even if declassified.
[43] The students fixed effects make it unnecessary to include time invariant demographic and grade of classification indicators on their own.

VI. Results

Baseline and ITT Models

As shown in Table 3, both our baseline and ITT models suggest a positive effect of special education for LDs. Baseline models with school fixed effects yield positive and statistically significant effects of .116 and .055, in math and ELA respectively, and of .326 percentage points in attendance (about a half day of school). Point estimates are smaller in the baseline student fixed effects models – .071 (Math), .063 (ELA), and .282 (attendance). The intent to treat (ITT) estimates of impacts on performance are larger in both school and student fixed effects models, although the attendance effect shrinks. In our preferred ITT student fixed effects specification, academic effects are both statistically and substantively significant at .117 in math and .102 in ELA, at or above the 0.1 effect size criterion suggested by Bloom *et al.* (2006) for a successful educational intervention, while the attendance effect shrinks considerably (.063 p.p.).

Heterogeneity by Gender and Race

As shown in Table 4, academic effects differ by gender and race. In this specification, Hispanic boys are the reference group and we estimate effects of .068 and .080 in math and ELA, respectively, somewhat smaller than the average effects reported above. Impacts are larger for girls, by .059 in Math and .017 in ELA, although only the math coefficient is statistically significant. Effects for Asians are much larger in math (.097) with no significant difference in ELA while smaller for blacks in both subjects (-.057 and -.068 in math and ELA, respectively) and nearly eroding the main effects (of .068 and .080). Effects for whites are also smaller than for Hispanics, but the difference is only significant (substantively and statistically) in ELA. As for whites, the white interaction effect for ELA is so large (-.074) that the net effect on whites is practically zero. There is less heterogeneity in the attendance effects – none suggests large effects of even as much as a day – although there is a statistically significant positive effect on black students and a negative effect on white students. In sum, these results suggest a positive impact of special education on test scores for all students, with the largest effects for Asians and girls in math.

Heterogeneity by Grade of First Classification

As shown in Table 5, estimated impacts also vary by grade of classification – with the largest effects among students classified in the earlier grades. Specifically, the largest, positive effects are for students classified in grades 4 (LD4) and 5 (LD5): .158 and .147 in math and .171 and .093 in ELA, respectively. For those classified in grade 6 (LD6), the effect in math is statistically significant (.066), but insignificant in ELA (.013). Finally, there is no statistically significant effect in either subject for students classified in grade 7 (LD7). These larger effects for early classification are consistent with hypotheses described earlier - the particular benefits of supports in earlier grades, duration of receipt, or efficacy of parent or school advocacy in earlier grades – although we cannot distinguish between these.

We estimate attendance models including the wider set of students classified between grades 1 and 7 (LD1 through LD7), finding a similar pattern of largest effects in the earlier classification cohorts followed by steady declines. Specifically, we find positive and statistically significant, albeit small, effects for students in early grades 1 through 5 – slightly less than one additional day per year. For those classified late, we find negative effects of -.309 and -1.354 for LD6 and LD7, respectively.

Pre- and Post- Classification Outcomes: Event Study

We further explore the effects of special education by tracing the time-path of outcomes pre- and post-classification in an event study framework, separately by classification cohort for two reasons. First, we make transparent differences in the length of pre- and post-classification periods for each classification cohort, and second, we explore differences by grade of classification. To be concrete, for LD4, we have only one pre- and four post- periods; for LD5, we have two pre- and three post-; for LD6, we have three pre- and three post; and for LD7, we have four pre- and only one post-. Thus, we are able to estimate more of the post-classification time-path for the earlier cohorts and more of the pre-classification time-path for the later cohorts.

As shown in Table 6, we find consistent evidence of higher performance two to three years prior to classification with a pattern of dropping performance for the later cohorts. More specifically, LD5 drops .064 and .110, LD6 drops .102 and .173, and LD7 drops .228 and .334 in math and ELA, respectively, between 3rd grade and classification.[44] As discussed earlier, a pre-classification pattern of dropping test scores is consistent with the hypotheses that parents, teachers, and/or school personnel respond to *dropping* (rather than just low) performance by considering referral for evaluation for special education classification.

Although it is common to regard pre-trends as a problem for identifying and interpreting treatment effects, this often reflects a concern that estimated treatment effects merely capture the impact of the pre-existing upward trends continuing post treatment - and estimates would be upwardly biased. In this case, we regard the pre-classification trend as a *feature* of the process of identifying students with disabilities for special education services in which case our estimation strategy is likely to yield *downwardly biased* impact estimates – if performance would have

[44] Concerning attendance (column 3), we see little evidence of any consistent trend pre-classification: some LDs attend school more often prior to classification (LD5), while others attend less (LD6 and LD7).

continued to drop in the absence of classification. Here, we see test scores *dropping* in the years prior to classification, as LD students find it harder and harder to keep up with the increasing academic demands without supports, and inducing parents or teachers to refer for evaluation. That said, one might be concerned that idiosyncratic declines in performance drive classification into special education, so that the subsequent mean reversion in test scores will upwardly bias estimated treatment effects (an Ashenfelter dip). We examine this possibility empirically by estimating models predicting LD classification in 5th, 6th, and 7th grade for all NYC public school GENs with annual changes in math or ELA achievement after 3rd grade (see Appendix Table A1). We find no substantively significant evidence that a one period drop in achievement predicts entry into special education with an LD classification.

Turning to the post-classification period, we find consistent evidence that outcomes improve after classification. More specifically, LD4, LD5, and LD7 all see significant increases in math and ELA performance immediately following classification (LD4 and LD5 show the largest increases of .070 and .138 in math and .096 and .079 in ELA, respectively, with LD7 showing increases of roughly .050 in both subjects). In subsequent periods, performance for LD4 shows some signs of fading – both in mathematics and ELA – particularly four years after classification in grade 8, which may (or may not) relate to the completion of middle school and/or preparation for entry into high school, which we discuss below. We see similar patterns for the other cohorts – positive effects in the short run and fading (or turning negative for ELA) in 8th grade.[45] We find little consistent pattern of attendance post-classification.

One explanation for the fading results could be that our intent to treat effects include the negative effects of declassification of students (out of special education and into general

[45] Since testing ends in grade 8, there are fewer "post" observations for students first classified in grades 6 and 7.

education) in later grades. To explore this, we extend the event study models and estimate separate coefficients in each post-classification year for LDs classified in special education in that year (*LDC*) and those who were declassified in that year (*GEN*). As shown in Table 7, in nearly all periods post-classification, declassified LDs perform worse than those still classified, although these differences are only statistically significant among students in grade 8 (e.g. *Post-4* math and ELA for LD4, *Post-3* math and ELA for LD5, and *Post-2* ELA for LD6).[46] Whether this poorer performance reflects selection out of special education – that is, students who would perform poorly are disproportionately declassified – or the negative effects of declassification itself is unclear, and we are unable to disentangle.

Probing the Results – Heterogeneity of Services

Finally, we explore the extent to which service settings differ across cohorts and time, estimating a series of event study style models with dependent variables capturing the different service settings.[47] We are particularly interested in addressing two key questions. First, does the initial service setting following classification differ across students classified at different grades? Second, do the trajectories differ across cohorts – that is, does the propensity to change service or be declassified differ by classification cohort?

As shown in Table 8, roughly two thirds of all LDs are placed in SETSS immediately following classification, with another 28 percent in ICT and a small share (5 percent) in SC classrooms. Thus, most LDs are initially placed in less restrictive settings (SETSS or ICT). There are, however, some differences in initial placements across classification cohorts. Students classified in grades 4 or 5 are more likely to be placed in SETSS (LD4, 68.1; LD5, 71.6) while

[46] There is no statistical difference in achievement between groups in early post-classification periods (e.g. *Post-1* for all subjects/classification cohorts, *Post-2* math for all cohorts, *Post-2* ELA for LD5, and *Post-3* for LD4 and LD5).

[47] These models do not include other controls and excludes a constant term to facilitate interpretation.

later cohorts are slightly more likely to begin in SC classes (LD6, 6.1; LD7, 7.4). The greater reliance on less restrictive placements (esp. SETSS) for earlier cohorts persists as student matriculate to higher grade levels. At the same time, there are differences in the timing of declassification - LD6 and LD7 are more likely to be declassified (15.0 and 19.0, respectively) after only one year (also statistically significant). Taken together, larger impacts of special education for LD4 and LD5 may be a result of initial placement in, and/or longer exposure to, less restrictive services (i.e. ICT and SETSS).

Robustness Checks

To test the robustness of our estimates, we re-estimate the baseline and all ITT models (including heterogeneity by gender, race/ethnicity, and grade of classification) under a set of alternative models/samples. First, we test the sensitivity of our specification by re-estimating without the controls for student mobility and grade retention (Table A2). Second, we re-estimate with a larger sample that includes all LDs with a secondary disability classification (Table A3). Finally, we re-estimate with a smaller sample, limiting the sample to LDs with three years of consecutive enrollment in the NYC public school system (Table A4).

Excluding controls for student mobility and grade retention yields substantially similar results, as does limiting the analytic sample to continuously enrolled LDs. Expanding the sample to include LDs with a secondary disability classification produces slightly different impact estimates by gender and race. In our original analysis (Table 4), the effect of classification into special education is larger for Asians in math (.059), smaller for whites in ELA (-.074), and smaller but insignificant for whites in math, compared to Hispanics. The robustness check in Table A2 shows no significant differences in effects for Asians in math or whites in ELA, but a larger and statistically significant effect for whites in math (.090).

To boost confidence in a causal interpretation of our results, we explore a set of estimated "effects" generated by a randomly generated classification. To do so, we re-estimate the ITT model with an *LD* variable constructed through a randomly generated classification date. That is, we randomly generate a new grade of classification variable (between four and eight for the math and ELA analysis and between one and eight for the attendance analysis) that matches the distribution of our original grade of classification variable. We repeat this process 500 times to calculate a distribution of impacts for both the ITT model and specifications that examine impacts by grade of classification. As shown in Appendix Table A5, estimated impacts are close to zero – that is, randomly assigned start dates yield very small and statistically insignificant impact estimates – thus, bolstering our confidence that our ITT estimates are not merely reflecting upward trends in the data.

VII. Conclusion

While the explosion of research on educational effectiveness drawing on administrative data and quasi-experimental (and experimental) methods over the last decades has yielded a bumper crop of evidence to guide policy for general education students, the bounty for special education has been thin. There is, instead, a dearth of rigorous research gauging the effectiveness of special education which sorely limits the ability of policymakers and educators to implement evidence based policy to improve outcomes for students with disabilities. The large population of students with disabilities and high cost of special education makes this gap particularly troubling. We build upon an important paper in the limited literature on special education, Hanushek, Rivkin, and Kain (2002), which studied Texas students in grades 4 to 7 and reported effects of .11 for LDs who remain in special education. We leverage rich, longitudinal data on K-8 students with disabilities in America's largest school district, NYC, to estimate the impact of special education on students classified with a learning disability after school entry. Using within-student pre/post comparisons and student fixed effects, we estimate impacts by grade of classification and explore differences across socio-demographic groups.

Our results indicate that special education "works" to improve outcomes for students with learning disabilities. We find substantively and statistically significant improvements in test scores of more than 0.1 standard deviation in both math and ELA over all LDs, with larger effects among those first classified in grades 4 or 5 than in later grades. The magnitude is substantively important, representing roughly 18% and 16% of the mean LD-GEN achievement gaps in math and ELA, respectively. Finally, our estimated effects are of similar magnitude to the Hanushek, Rivkin, and Kain (2002) findings for Texas, suggesting results may be generalizable to other contexts.

Turning to mechanisms, our exploratory models suggest inclusive part-time services -- Special Education Teacher Support Services and Related Services - may be particularly effective for LDs and, perhaps, expanding these placements could improve outcomes. That said, since placement into specific services is undoubtedly endogenous, at least to some extent, our estimates cannot be viewed as credibly causal. Instead, our descriptive findings suggest more work investigating the differential effects of alternative service settings could be useful for policymakers.

Finally, the particularly strong effects for girls and Asians and relatively poor results for black boys suggest further work is warranted, particularly in light of the uneven incidence of special education across demographic groups. The relatively high incidence of special education among black boys, compared to whites and girls, and the relatively small impact may bolster concerns about disparate impact and racism driving over-classification, in contrast to the

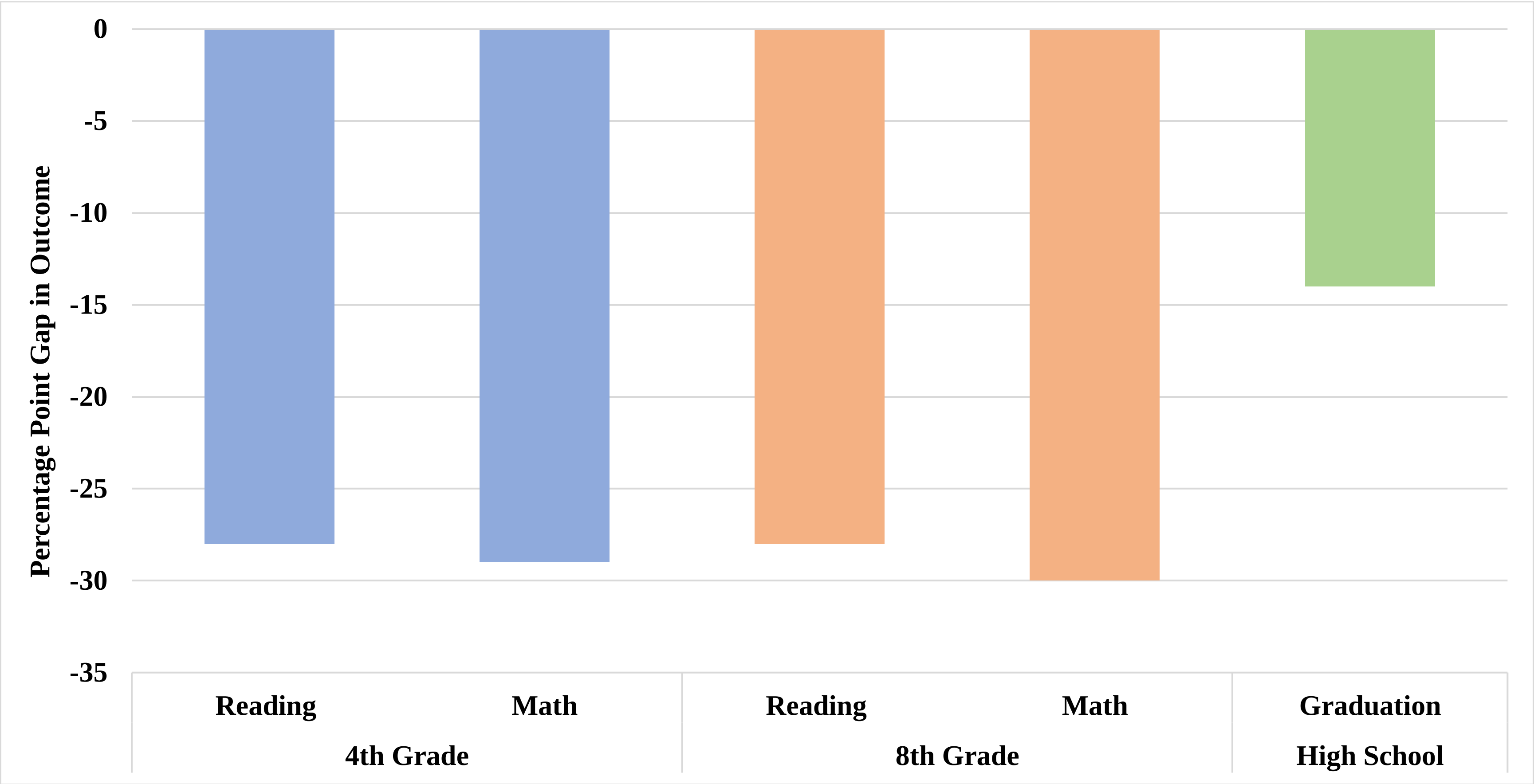
Morgan, *et al.* (2015) suggestion that blacks are *under*-represented relative to their underlying disabilities. Conversely, the relatively low incidence of special education among Asians and particularly large impact of special education may bolster concerns that Asian parents resist classification into special education such that only those Asian LDs with the most severe disabilities – perhaps, those with the greatest potential to benefit from services – are classified into special education (see Cooc, 2018).

To some extent, whether demographic differences in representation are problematic hinges on whether special education improves outcomes for enrolled students. If it works poorly, then disproportionality may be more readily viewed as problematic; if it works well, perceptions about disproportionality may be more positive. Our findings of positive effects overall may alleviate concerns for some, but the differences in impacts across demographic groups warrant further attention.

Overall, our work documents a meaningful positive effect of special education for students with learning disabilities and reveals provocative heterogeneity in the treatment effects. Finally, this paper demonstrates both the possibility of obtaining the requisite data and of leveraging quasi-experimental methods to derive credibly causal estimates of the efficacy of educational policy and practices and building the evidence base to improve special education for all.

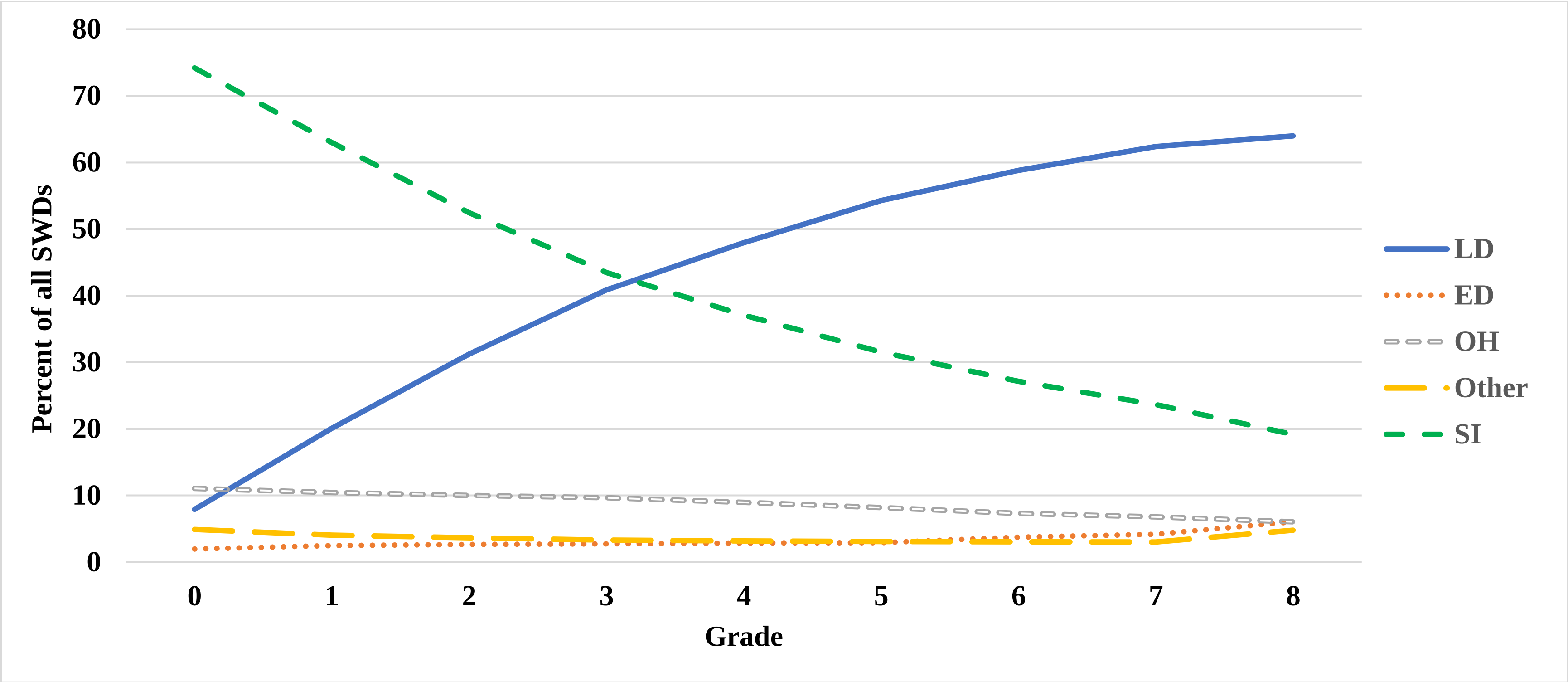
Figures and Tables

***Figure 1*:** SWD-GEN Achievement Gaps: 2017 NAEP Math and Reading, Grades 4 and 8; 2016 Public High School Graduation Rates



Sources: NAEP - 2017 Math & Reading Assessments. National Center for Education Statistics, 2018. The Condition of Education - Preprimary, Elementary, and Secondary Education – Elementary and Secondary Enrollment - Children and Youth with Disabilities, (2018). Retrieved from https://nces.ed.gov/programs/coe/indicator_cgg.asp.

***Figure 2*:** NYC SWD Classification by Grade, 2006-2012



Notes: SWD is identified by students who have an Individualized Education Program (IEP). In each year, approximately 21% of all LDs are classified as ungraded.

***Table 1*: Descriptive Statistics, NYC Public School Students, Grades K-8, 2006**

	Never SWD (GEN)	Ever SWD (non-LD)	Ever LD (only LD)
Demographics			
Female	51.9	28.6	41.3
Asian/Other	15.7	6.6	6.3
Black	31.0	34.0	34.8
Hispanic	38.7	44.3	46.4
White	14.7	15.0	12.5
Foreign Born	15.6	4.9	9.4
FRPL	76.7	81.7	84.7
ELL	11.5	16.7	19.5
Academic Outcomes			
Attendance (K-8)	91.6	89.2	89.5
Math Z-Scores (3-8)	0.100	-0.292	-0.669
ELA Z-Scores (3-8)	0.112	-0.412	-0.749
Disabilities			
ED	--	15.0	--
SI	--	57.2	--
OH	--	10.8	--
LD	--	--	100
Other	--	17.0	--
Services			
SC	--	45.1	23.0
ICT	--	15.6	16.7
SETSS	--	39.3	60.3
N-students	575,595	67,421	55,948

Notes: GENs were never classified with a disability between 2006 and 2012; non-LDs were classified with a disability other than LD between 2006 and 2012. LDs were classified with one disability, LD, between 2006 and 2012. Excluding ungraded students, the special education population makes up 12.3% of the student body in 2006. FRPL and ELL denote students eligible for free or reduced price lunch and English language learner services, respectively. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one.

Table 2: Characteristics of LDs by Grade of Classification, LDs classified between 3rd and 8th Grade, 2006-2012

	4th	5th	6th	7th
Demographics				
Female	46.1	47.9	44.0	40.2
Asian/Other	10.1	8.9	6.5	5.3
Black	31.9	32.7	35.7	38.6
Hispanic	45.8	47.3	48.2	48.1
White	12.2	11.1	9.5	7.9
Foreign Born	11.0	12.8	14.4	14.0
FRPL	86.8	86.0	86.0	86.3
ELL	21.1	21.5	21.1	16.8
Initial Service				
SC	3.8	3.0	6.0	7.3
ICT	28.0	25.5	32.1	27.2
SETSS	68.2	71.3	60.7	64.9
N-students	6,309	5,471	3,864	3,604
Academic Outcomes in Grade 3				
Attendance	92.4	92.0	90.7	88.7
ELA Z-Score	-0.934	-0.796	-0.751	-0.700
Math Z-Score	-0.868	-0.836	-0.839	-0.766
N-students	6,288	5,179	3,451	3,108

Notes: The sample includes SWDs classified in 3rd grade and after with an LD only. FRPL and ELL denote students eligible for free or reduced price lunch and English language learner services, respectively. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one.

Table 3: Effects of Special Education on Academic Performance, Grades 3-8 for Math and ELA, Grades K-8 for Attendance, 2006 –2012

	Math		ELA		Attendance	
	(1)	(2)	(3)	(4)	(5)	(6)
Baseline						
LDC	0.116***	0.071***	0.055***	0.063***	0.326***	0.282***
	(0.009)	(0.005)	(0.008)	(0.006)	(0.055)	(0.049)
R^2	0.252	0.746	0.259	0.742	0.150	0.720
Intent to Treat						
LD	0.360***	0.117***	0.285***	0.102***	0.296***	0.068
	(0.009)	(0.007)	(0.008)	(0.008)	(0.057)	(0.051)
R^2	0.274	0.747	0.275	0.742	0.150	0.720
School FE	Y	Y	Y	Y	Y	Y
Student FE	N	Y	N	Y	N	Y
N-observations	92,902	92,902	90,356	90,356	197,274	197,274
N-students	24,189	24,189	23,901	23,901	44,487	44,487

Standard errors in parentheses. ** $p < 0.05$, *** $p < 0.01$

Notes: All models include time-varying student demographic characteristics (i.e. *FRPL, ELL, Mobility*, and *HeldBack*) and grade, year, school, and student fixed effects**.** The sample includes SWDs classified with an LD only. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one. Attendance is measure on a scale from 0 to 100.

Table 4: Effects of Special Education by Demographics, Grades 3-8 for Math and ELA Z-scores, Grades K-8 for Attendance, 2006 –2012

	Math (1)	ELA (2)	Attendance (3)
LD	0.068***	0.080***	0.033
	(0.011)	(0.011)	(0.077)
Female*LD	0.059***	0.017	-0.029
	(0.010)	(0.010)	(0.072)
Asian*LD	0.097***	0.025	0.210
	(0.023)	(0.024)	(0.125)
Black*LD	-0.057***	-0.068***	0.319***
	(0.013)	(0.012)	(0.098)
White*LD	-0.025	-0.074***	-0.221**
	(0.018)	(0.017)	(0.104)
School FE	Y	Y	Y
Student FE	Y	Y	Y
N-observations	92,902	90,356	197,274
N-students	24,189	23,901	44,487
R^2	0.749	0.745	0.720

Standard errors in parentheses. ** $p < 0.05$, *** $p < 0.01$

Notes: All models include time-varying student demographic characteristics (i.e. *FRPL, ELL, Mobility*, and *HeldBack*) and grade, year, school, and student fixed effects**.** The sample includes SWDs classified with an LD only. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one. Attendance is measure on a scale from 0 to 100. Hispanic boys are the omitted (comparison) group.

Table 5: Effects of Special Education on Academic Performance by Grade of Classification, Grades 3-8 for Math and ELA, Grades K-8 for Attendance, 2006 –2012

	Math (1)	ELA (2)	Attendance (3)
LD1			0.874***
			(0.146)
LD2			0.299***
			(0.092)
LD3			0.228***
			(0.081)
LD4	0.158***	0.171***	0.312***
	(0.013)	(0.013)	(0.085)
LD5	0.147***	0.093***	0.213**
	(0.011)	(0.012)	(0.104)
LD6	0.066***	0.013	-0.309**
	(0.012)	(0.013)	(0.154)
LD7	-0.005	-0.024	-1.354***
	(0.015)	(0.015)	(0.216)
School FE	Y	Y	Y
Student FE	Y	Y	Y
N-observations	92,902	90,356	197,274
N-students	24,189	23,901	44,487
F statistic	34.72***	38.65***	13.46***
R^2	0.747	0.743	0.726

Standard errors in parentheses. ** $p < 0.05$, *** $p < 0.01$

Notes: All models include time-varying student demographic characteristics (i.e. *FRPL*, *ELL*, *Mobility*, and *HeldBack*) and grade, year, school, and student fixed effects**.** The sample includes SWDs classified with an LD only. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one. Attendance is measure on a scale from 0 to 100. The asterisks on the *F* statistics indicate the level of significance at which the null hypothesis – there is no difference in estimates by grade of classification – can be rejected.

Table 6: Event Study: Academic Performance, Grades 3-8 for Math, ELA, and Attendance, 2006-2012

		Math (1)	ELA (2)	Attendance (3)
LD4	Pre-1 (omitted)			
	Classification	0.070***	0.096***	0.394***
		(0.016)	(0.017)	(0.139)
	Post-1	0.116***	0.084***	0.326**
		(0.020)	(0.020)	(0.165)
	Post-2	0.101***	0.067***	0.358
		(0.021)	(0.020)	(0.221)
	Post-3	0.095***	0.062***	0.899***
		(0.022)	(0.022)	(0.249)
	Post-4	0.044	-0.038	0.190
		(0.028)	(0.026)	(0.326)
LD5	Pre-2	0.064***	0.110***	-0.531***
		(0.016)	(0.018)	(0.161)
	Pre-1 (omitted)			
	Classification	0.138***	0.079***	0.269
		(0.016)	(0.018)	(0.146)
	Post-1	0.134***	0.058***	0.084
		(0.019)	(0.019)	(0.219)
	Post-2	0.133***	0.050**	0.561**
		(0.020)	(0.021)	(0.255)
	Post-3	0.113***	-0.044	-0.446
		(0.028)	(0.024)	(0.343)
LD6	Pre-3	0.102***	0.173***	-0.059
		(0.025)	(0.025)	(0.253)
	Pre-2	0.023	0.069***	0.184
		(0.020)	(0.023)	(0.194)
	Pre-1 (omitted)			
	Classification	0.033	-0.001	0.136
		(0.020)	(0.020)	(0.241)
	Post-1	0.074***	0.019	0.432
		(0.023)	(0.022)	(0.307)
	Post-2	0.052	-0.096***	-0.977**
		(0.029)	(0.026)	(0.397)
LD7	Pre-4	0.228***	0.334***	1.353***
		(0.029)	(0.029)	(0.362)
	Pre-3	0.171***	0.268***	1.132***
		(0.028)	(0.024)	(0.329)
	Pre-2	0.126***	0.143***	0.505*
		(0.022)	(0.023)	(0.260)
	Pre-1 (omitted)			
	Classification	0.054**	0.051**	0.207
		(0.021)	(0.022)	(0.291)
	Post-1	0.037	-0.053**	-1.623***
		(0.030)	(0.027)	(0.413)
School FE		Y	Y	Y
Student FE		Y	Y	Y
N-observations		92,902	90,356	197,274
N-students		24,189	23,901	44,487
R^2		0.749	0.745	0.745

Standard errors in parentheses. ** $p < 0.05$, *** $p < 0.01$

Notes: All models include time-varying student demographic characteristics (i.e. *FRPL*, *ELL*, *Mobility*, and *HeldBack*) and grade, year, school, and student fixed effects**.** The sample includes SWDs classified with an LD only. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one. Attendance is measure on a scale from 0 to 100.

***Table 7*: Event Study: Academic Performance Separating Post-Classification Trends by Special Education Status, Grades 3-8 for Math, ELA, and Attendance, 2006-2012**

			Math (1)	ELA (2)	Attendance (3)
LD4	Pre-1 (omitted)				
	Classification		0.069^{***}	0.095^{***}	0.372^{***}
			(0.016)	(0.017)	(0.138)
	Post-1	LDC	0.116^{***}	0.084^{***}	0.321
			(0.020)	(0.020)	(0.164)
		GEN	0.102^{**}	0.051	-0.258
			(0.046)	(0.043)	(0.467)
	Post-2	LDC	0.101^{***}	0.073^{***}	0.485^{**}
			(0.021)	(0.021)	(0.214)
		GEN	0.095^{***}	0.004	-1.114
			(0.033)	(0.035)	(0.568)
	Post-3	LDC	0.101^{***}	0.066^{***}	0.955^{***}
			(0.023)	(0.023)	(0.258)
		GEN	0.064^{**}	0.033	0.422
			(0.029)	(0.033)	(0.362)
	Post-4	LDC	0.059^{**}	-0.024	0.131
			(0.030)	(0.026)	(0.357)
		GEN	0.004	-0.078^{**}	0.246
			(0.035)	(0.034)	(0.412)
LD5	Pre-2		0.063^{***}	0.110^{***}	-0.534^{***}
			(0.016)	(0.018)	(0.161)
	Pre-1 (omitted)				
	Classification		0.138^{***}	0.078^{***}	0.176
			(0.016)	(0.018)	(0.145)
	Post-1	LDC	0.132^{***}	0.059^{***}	0.180
			(0.019)	(0.019)	(0.215)
		GEN	0.136^{***}	0.036	-1.947^{***}
			(0.038)	(0.035)	(0.604)
	Post-2	LDC	0.133^{***}	0.051^{**}	0.709^{***}
			(0.020)	(0.021)	(0.250)
		GEN	0.120^{***}	0.028	-1.208^{**}
			(0.036)	(0.037)	(0.587)
	Post-3	LDC	0.121^{***}	-0.038	-0.389
			(0.029)	(0.024)	(0.353)
		GEN	0.075	-0.074^{**}	-0.958
			(0.040)	(0.036)	(0.497)
LD6	Pre-3		0.101^{***}	0.173^{***}	-0.046
			(0.025)	(0.025)	(0.253)
	Pre-2		0.022	0.069^{***}	0.208
			(0.020)	(0.023)	(0.194)
	Pre-1 (omitted)				
	Classification		0.032	-0.004	0.041
			(0.020)	(0.020)	(0.242)
	Post-1	LDC	0.073^{***}	0.018	0.506
			(0.023)	(0.022)	(0.300)
		GEN	0.067	-0.001	-0.914
			(0.038)	(0.038)	(0.791)
	Post-2	LDC	0.056	-0.085^{***}	-0.789^{**}
			(0.030)	(0.027)	(0.398)
		GEN	0.021	-0.168^{***}	-2.158^{***}
			(0.045)	(0.040)	(0.806)
LD7	Pre-4		0.228^{***}	0.333^{***}	1.339^{***}
			(0.029)	(0.029)	(0.361)
	Pre-3		0.171^{***}	0.268^{***}	1.144^{***}
			(0.028)	(0.024)	(0.328)
	Pre-2		0.126^{***}	0.143^{***}	0.518^{**}
			(0.022)	(0.023)	(0.258)
	Pre-1 (omitted)				
	Classification		0.052^{**}	0.052^{**}	-0.042
			(0.022)	(0.022)	(0.291)
	Post-1	LDC	0.040	-0.056^{**}	-1.208^{***}
			(0.030)	(0.027)	(0.392)
		GEN	0.004	-0.032	-4.813^{***}
			(0.058)	(0.048)	(1.130)
School FE			Y	Y	Y
Student FE			Y	Y	Y
N-observations			92,902	90,356	197,274
N-students			24,189	23,901	44,487
R^2			0.749	0.745	0.746

Standard errors in parentheses. ** $p < 0.05$, *** $p < 0.01$

Notes: All models include time-varying student demographic characteristics (i.e. *FRPL, ELL, Mobility*, and *HeldBack*) and grade, year, school, and student fixed effects. The sample includes SWDs classified with an LD only. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one. Attendance is measure on a scale from 0 to 100.

Table 8: Event Study: Service, Grades 3-8 for Math and ELA, Grades K-8 for Attendance, 2006 –2012

		SC (1)	ICT (2)	SETSS (3)	GEN (4)
LD4	Pre-1				1.000***
					(0.000)
	Classification	0.038***	0.280***	0.681***	
		(0.003)	(0.011)	(0.011)	
	Post-1	0.005***	0.293***	0.629***	0.073***
		(0.001)	(0.012)	(0.012)	(0.005)
	Post-2	0.012***	0.309***	0.568***	0.111***
		(0.002)	(0.011)	(0.011)	(0.006)
	Post-3	0.013***	0.305***	0.485***	0.197***
		(0.002)	(0.012)	(0.013)	(0.010)
	Post-4	0.008***	0.262***	0.422***	0.308***
		(0.002)	(0.013)	(0.014)	(0.013)
LD5	Pre-1				1.000***
					(0.000)
	Classification	0.029***	0.255***	0.716***	
		(0.003)	(0.011)	(0.011)	
	Post-1	0.007***	0.268***	0.620***	0.105***
		(0.001)	(0.010)	(0.011)	(0.007)
	Post-2	0.009***	0.266***	0.585***	0.139***
		(0.002)	(0.011)	(0.013)	(0.009)
	Post-3	0.010***	0.255***	0.497***	0.238***
		(0.002)	(0.013)	(0.015)	(0.013)
LD6	Pre-1				1.000***
					(0.000)
	Classification	0.061***	0.320***	0.619***	
		(0.005)	(0.012)	(0.013)	
	Post-1	0.012***	0.314***	0.524***	0.150***
		(0.002)	(0.014)	(0.014)	(0.011)
	Post-2	0.027***	0.286***	0.476***	0.211***
		(0.004)	(0.015)	(0.016)	(0.013)
LD7	Pre-1				1.000***
					(0.000)
	Classification	0.074***	0.273***	0.652***	
		(0.005)	(0.013)	(0.014)	
	Post-1	0.028***	0.260***	0.523***	0.189***
		(0.004)	(0.014)	(0.017)	(0.012)
N-observations		95,635	95,635	95,635	95,635
N-students		24,390	24,390	24,390	24,390
R^2		0.025	0.221	0.471	0.578

Standard errors in parentheses. ** $p < 0.05$, *** $p < 0.01$

Notes: Models contain no additional controls or no constant term. Coefficients times 100 indicate the percent of LDs by type of service.

Appendix

Table A1: Predicting Entry into Special Education with LD Diagnosis, All NYC Public School Students, Grades 4-7, 2006-2012

	Classified as LD in 5th Grade (1)	Classified as LD in 6th Grade (2)	Classified as LD in 7th Grade (3)
Male	0.002***	0.002***	0.003***
	(0.000)	(0.000)	(0.000)
Black	-0.004***	0.000	0.002***
	(0.001)	(0.000)	(0.001)
Hispanic	-0.003***	0.000	0.002***
	(0.001)	(0.000)	(0.001)
Asian/Other	-0.002***	-0.001***	-0.001**
	(0.001)	(0.000)	(0.000)
Native Born	0.003***	0.003***	0.003***
	(0.001)	(0.000)	(0.000)
ELL	0.008***	0.008***	0.004***
	(0.001)	(0.001)	(0.001)
FRPL	-0.002***	-0.001	-0.000
	(0.001)	(0.000)	(0.000)
3rd Grade ELA	-0.008***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)
3rd Grade Math	-0.007***	-0.006***	-0.004***
	(0.000)	(0.000)	(0.000)
3rd-4th Change ELA	0.007***	0.002***	0.001***
	(0.000)	(0.000)	(0.000)
3rd-4th Change Math	0.005***	0.005***	0.003***
	(0.000)	(0.000)	(0.000)
4th-5th Change ELA		-0.001***	0.000
		(0.000)	(0.000)
4th-5th Change Math		0.002***	0.002***
		(0.000)	(0.000)
5th-6th Change ELA			0.000
			(0.000)
5th-6th Change Math			0.002***
			(0.000)
Year FE	Y	Y	Y
N-students	330,752	324,610	328,331
R^2	0.012	0.006	0.004

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

Notes: Each column uses student characteristics of general education students to predict LD classification in the following grade. E.g. Column 1 presents the relationship between student characteristics of general education students in 4th grade and classification as LD in the 5th grade.

Table A2: Effects of Special Education on Academic Performance, Grades 3-8 for Math and ELA, Grades K-8 for Attendance, 2006-2012, *Alternate Specification -No Controls for Mobility* and *Grade Retention*

	Math				ELA				Attendance			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
LDC	0.071***				0.062***				0.301***			
	(0.006)				(0.006)				(0.049)			
LD		0.118***	0.072***			0.102***	0.080***			0.078	0.032	
		(0.007)	(0.011)			(0.008)	(0.011)			(0.051)	(0.077)	
Female			0.059***				0.016				-0.039	
			(0.010)				(0.010)				(0.072)	
Asian			0.097***				0.043				0.156	
			(0.022)				(0.024)				(0.124)	
Black			-0.067***				-0.072***				0.394***	
			(0.013)				(0.012)				(0.098)	
White			-0.026				-0.077***				-0.297***	
			(0.018)				(0.017)				(0.104)	
LD1												0.846***
												(0.145)
LD2												0.284***
												(0.091)
LD3												0.230***
												(0.081)
LD4				0.162***				0.171***				0.324***
				(0.013)				(0.013)				(0.085)
LD5				0.151***				0.094***				0.237**
				(0.011)				(0.012)				(0.104)
LD6				0.073***				0.016				-0.257*
				(0.012)				(0.013)				(0.154)
LD7				-0.011				-0.026				-1.351***
				(0.015)				(0.015)				(0.215)
N-observations	92,902	92,902	92,902	92,902	90,356	90,356	90,356	90,356	197,274	197,274	197,274	197,274
N-students	24,189	24,189	24,189	24,189	23,901	23,901	23,901	23,901	44,487	44,487	44,487	44,487
R^2	0.745	0.745	0.748	0.746	0.742	0.742	0.745	0.743	0.720	0.720	0.720	0.720

Standard errors in parentheses. ** $p < 0.05$, *** $p < 0.01$

Notes: All models include time-varying student demographic characteristics (i.e. *FRPL* and *ELL*) and grade, year, school, and student fixed effects**.** Sample includes SWDs classified with only an LD. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one. Attendance is measure on a scale from 0 to 100.

Table A3: Effects of Special Education on Academic Performance, Grades 3-8 for Math and ELA, Grades K-8 for Attendance, 2006-2012, *Alternate Sample: LDs with Secondary Disability*

	Math				ELA				Attendance			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
LDC	0.064***				0.057***				0.306***			
	(0.005)				(0.005)				(0.046)			
LD		0.113***	0.066***			0.098***	0.079***			0.072	-0.000	
		(0.007)	(0.011)			(0.007)	(0.011)			(0.049)	(0.073)	
Female			0.059***				0.017				0.004	
			(0.010)				(0.010)				(0.069)	
Asian			-0.030				-0.085***				0.180	
			(0.017)				(0.016)				(0.124)	
Black			-0.057***				-0.071***				0.422***	
			(0.013)				(0.012)				(0.095)	
White			0.090***				0.036				-0.322***	
			(0.021)				(0.023)				(0.099)	
LD1												0.880***
												(0.132)
LD2												0.316***
												(0.087)
LD3												0.255***
												(0.078)
LD4				0.148***				0.159***				0.281***
				(0.013)				(0.012)				(0.082)
LD5				0.138***				0.085***				0.126
				(0.011)				(0.012)				(0.104)
LD6				0.063***				0.007				-0.343**
				(0.012)				(0.013)				(0.153)
LD7				-0.007				-0.023				-1.428***
				(0.015)				(0.015)				(0.217)
N-observations	100,725	100,725	100,725	100,725	98,058	98,058	98,058	98,058	218,609	218,609	218,609	218,609
N-students	26,325	26,325	26,325	26,325	25,831	25,831	25,831	25,831	49,098	49,098	49,098	49,098
R^2	0.748	0.748	0.750	0.748	0.740	0.740	0.743	0.741	0.718	0.718	0.718	0.718

Standard errors in parentheses. ** $p < 0.05$, *** $p < 0.01$

Notes: All models include time-varying student demographic characteristics (i.e. *FRPL* and *ELL*) and grade, year, school, and student fixed effects**.** Sample includes LDs with a secondary disability classification. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one. Attendance is measure on a scale from 0 to 100.

Table A4: Effects of Special Education on Academic Performance, Grades 3-8 for Math and ELA, Grades K-8 for Attendance, 2006-2012, *Alternate Sample – 3 Years Consecutive Enrollment*

	Math				ELA				Attendance			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
LDC	0.064***				0.057***				0.248***			
	(0.006)				(0.006)				(0.049)			
LD		0.108***	0.063***			0.091***	0.073***			0.080	0.051	
		(0.007)	(0.011)			(0.008)	(0.011)			(0.050)	(0.076)	
Female			0.057***				0.019				-0.039	
			(0.010)				(0.010)				(0.071)	
Asian			0.100***				0.035				0.106	
			(0.022)				(0.022)				(0.121)	
Black			-0.058***				-0.072***				0.340***	
			(0.013)				(0.012)				(0.100)	
White			-0.019				-0.076***				-0.321***	
			(0.018)				(0.017)				(0.104)	
LD1												0.938***
												(0.152)
LD2												0.341***
												(0.090)
LD3												0.272***
												(0.080)
LD4				0.157***				0.168***				0.312***
				(0.014)				(0.014)				(0.084)
LD5				0.143***				0.090***				0.218**
				(0.011)				(0.012)				(0.102)
LD6				0.065***				0.015				-0.270
				(0.012)				(0.013)				(0.150)
LD7				-0.007				-0.025				-1.397***
				(0.015)				(0.015)				(0.215)
N-observations	81,790	81,790	81,790	81,790	79,966	79,966	79,966	79,966	180,400	180,400	180,400	180,400
N-students	18,633	18,633	18,633	18,633	18,706	18,706	18,706	18,706	36,050	36,050	36,050	36,050
R^2	0.718	0.719	0.721	0.719	0.701	0.702	0.705	0.702	0.702	0.702	0.702	0.703

Standard errors in parentheses. ** $p < 0.05$, *** $p < 0.01$

Notes: All models include time-varying student demographic characteristics (i.e. *FRPL, ELL, Mobility,* and *HeldBack*) and grade, year, school, and student fixed effects**.** Sample includes SWDs classified with only an LD enrolled in NYC public schools for at least 3 consecutive academic years. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one. Attendance is measure on a scale from 0 to 100

Table A5: Effects of Special Education on Academic Performance, Random Assignment of First Classification, Grades 3-8 for Math and ELA, 2006-2012

	Math			ELA			Attendance		
	Average Estimate		SD	Average Estimate		SD	Average Estimate		SD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LD	-0.000		0.0055	-0.000		0.0056	0.002		0.0051
	(0.006)		0.0002	(0.006)		0.0002	(0.005)		0.0001
LD4		0.001	0.0137		-0.000	0.0136		-0.000	0.0130
		(0.016)	0.0006		(0.015)	0.0006		(0.010)	0.0005
LD5		0.000	0.0104		0.001	0.0105		0.003	0.0109
		(0.012)	0.0004		(0.013)	0.0003		(0.011)	0.0004
LD6		-0.000	0.0118		0.000	0.0118		0.001	0.0111
		(0.010)	0.0003		(0.011)	0.0004		(0.011)	0.0005
LD7		0.000	0.0106		0.000	0.0105		0.002	0.0109
		(0.011)	0.0004		(0.012)	0.0004		(0.014)	0.0003
N-observations	92,902	92,902		90,356	90,356		197,274	197,274	
N-students	24,189	24,189		23,901	23,901		44,487	44,487	

Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Notes: All models include time-varying student demographic characteristics (i.e. *FRPL*, *ELL*, *Mobility*, and *HeldBack*) and grade, year, school, and student fixed effects**.** The sample includes SWDs classified with an LD only. Math and ELA Z-Scores are standardized for each grade, citywide, with a mean of zero and a standard deviation of one. Each model was estimated 500 times.

References

Author (2019)

Blackorby, J., & Wagner, M. (1996). Longitudinal postschool outcomes of youth with disabilities: Findings from the National Longitudinal Transition Study. *Exceptional children*, *62*(5).

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

The Condition of Education - Preprimary, Elementary, and Secondary Education - Elementary and Secondary Enrollment - Children and Youth with Disabilities. (2018). Retrieved from https://nces.ed.gov/programs/coe/indicator_cgg.asp

Cooc, N. (2018). Examining the Underrepresentation of Asian Americans in Special Education: New Trends from California School Districts. *Exceptionality*, *26*(1), 1-19.

Coutinho, M. & Oswald, D. (2000). Disproportionate representation in special education: A synthesis and recommendations. *Journal of Child and Family Studies*, 9, 135-156.

Cruz, R. A., & Rodl, J. E. (2018). An Integrative Synthesis of Literature on Disproportionality in Special Education. *The Journal of Special Education*, *52*(1), 50-63.

Fiester, L. (2010). Early Warning! Why Reading by the End of Third Grade Matters. KIDS COUNT Special Report. *Annie E. Casey Foundation*.

Ford, D. Y. (2008). Intelligence testing and cultural diversity: The need for alternative instruments, policies, and procedures. In J. L. Van Tassel-Baska (Ed.), Alternative Assessments with Gifted and Talented Students (pp. 107-128). Waco, TX: Prufrock Press.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Math instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, *79*(3), 1202-1242.

Glander, M. (2016). 2014–15 Common Core of Data (CCD) Universe Files (NCES 2016-077). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch/.

Guin, K. (2004). Chronic teacher turnover in urban elementary schools. Educational Evaluation and Policy Analysis, 12(42), 1–25.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2002). Inferring program effects for special populations: Does special education raise achievement for students with disabilities? *Review of Economics and Statistics*, *84*(4).

Hibel, J., Farkas, G., & Morgan, P. L. (2010). Who is placed into special education? *Sociology of Education*, *83*(4), 312-332.

Hocutt, A. M. (1996). Effectiveness of special education: Is placement the critical factor? *The future of children*, 77-102.

Individuals with Disabilities Education Act, 20 U.S.C. § 1400 (2004)

Morgan, P.L., Farkas, G., Hillemeier, M.M., Mattison, R., Maczuga, S., Li, H., & Cook, M. (2015). Minorities are Disproportionately Underrepresented in Special Education: Longitudinal Evidence Across Five Disability Conditions. *Educational Researcher*, 44(5), 278-292

Morgan, P. L., Frisco, M. L., Farkas, G., & Hibel, J. (2010). A propensity score matching analysis of the effects of special education services. *The Journal of special education*, *43*(4), 236-254.

Morgan, P. L., Sideridis, G., & Hua, Y. (2012). Initial and over-time effects of fluency interventions for students with or at risk for disabilities. *The Journal of Special Education*, *46*(2), 94-116.

Morgan, P. L., Farkas, G., Cook, M., Strassfeld, N. M., Hillemeier, M. M., Pun, W. H., & Schussler, D. L. (2017). Are Black children disproportionately overrepresented in special education? A best-evidence synthesis. *Exceptional Children, 83*(2), 181-198.

NYC Department of Education. (2018a). The IEP. Retrieved 2018, from https://www.schools.nyc.gov/special-education/the-iep-process/the-iep

NYC Department of Education. (2018b). Other Special Education Services. Retrieved 2018, from https://www.schools.nyc.gov/special-education/supports-and-services/other-special-education-services

National Center for Education Statistics. (2018). NAEP data explorer.

Parrish, T., Harr, J., Anthony, J., Merickel, A., & Esra, P. (2003). State Special Education Finance Systems, 1999-2000. Part I. *American Institutes for Research*.

Phelps, L. A., & Hanley-Maxwell, C. (1997). School-to-work transitions for youth with disabilities: A review of outcomes and practices. *Review of educational research*, *67*(2), 197-226.

Reynolds, A. J., & Wolfe, B. (1999). Special education and school achievement: An exploratory analysis with a central-city sample. *Educational evaluation and policy analysis*, *21*(3), 249-269.

Snyder, T. D. (1993). *120 years of American education: A statistical portrait*. DIANE Publishing.

Swanson, H. L. (1999). Reading research for students with LD: A meta-analysis of intervention outcomes. *Journal of learning disabilities*, *32*(6), 504-532.

Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject-design intervention research for students with LD. *Journal of learning disabilities*, *33*(2), 114-136.

Swanson, H. L., Hoskyn, M., & Lee, C. (1999). Interventions for students with learning disabilities: A meta-analysis of treatment outcomes. New York, NY, US: Guilford Press.

U.S. Department of Education, National Center for Education Statistics. (2018). *Digest of Education Statistics, 2016* (NCES 2017-094), Chapter 2.

Wagner, M., Newman, L., Cameto, R., Levine, P., & Garza, N. (2006). An Overview of Findings from Wave 2 of the National Longitudinal Transition Study-2 (NLTS2). NCSER 2006-3004. *National Center for Special Education Research.*