

# Are Two Teachers Better Than One?

## The Effect of Co-Teaching on Students With and Without Disabilities

Nathan Jones, PhD

Boston University

Marcus A. Winters, PhD

Boston University

### **Abstract**

Co-teaching, in which a general education teacher and special education teacher collaboratively instruct students with and without disabilities, has emerged as a common strategy for educating students within inclusive classrooms. We leverage longitudinal administrative data in Massachusetts to measure the causal effect of co-teaching on the test scores of students with and without disabilities. Co-teaching has a no effect or a small positive effect on math scores. In English language arts, co-teaching has a positive effect for students with disabilities in elementary grades, but a negative effect for both students with and without disabilities in secondary grades.

The data can be obtained by filing a request directly with the Massachusetts Department of Elementary and Secondary Education (MADESE). The authors are willing to assist. Neither author received financial support or has a conflicting interest. MADESE had the right to review the paper prior to circulation. We would like to thank Allison Gilmour, Tim Sass, participants at the APPAM annual conference, and participants at the Half Baked Seminar at The Annenberg Institute at Brown University for helpful comments, criticisms, and suggestions. We thank the Massachusetts Department of Elementary and Secondary Education for their cooperation.

JEL: I20; I24; I28

# I Introduction

About 13.2% of public school students in the United States receive special education services.<sup>1</sup>The proportion of students with disabilities (SWDs) spending at least 80% of instructional time in a general classroom setting has increased from 48% in 2001 to 63% today.<sup>2</sup> The motivations behind including SWDs in general education are at least as much social as they are academic. And yet, nearly 45 years since the signing of the federal Individuals With Disabilities Education Act (IDEA) that required schools to educate SWDs in the least restrictive environment, we have virtually no rigorous evidence on inclusion's impact on academic outcomes for both students with and without disabilities or about how to most effectively instruct students within inclusive classroom settings (Gilmour, 2019). Inclusive classroom settings are potentially challenging for general educators because they must adapt their teaching to meet the legally required instructional needs of SWDs (Friend, 2015).

Co-teaching, in which a general education teacher and special education teacher collaboratively provide instruction to students with and without disabilities in the same classroom (Friend, 2008), has emerged as a common strategy for educating SWDs within inclusive classroom settings. Co-teaching is meant to simultaneously provide SWDs access to the general education curriculum and the specially designed instruction outlined in their Individualized Education Programs (IEPs) (Friend et al., 2010; Friend, 2015). Further, with the collaborative work of two teachers, co-teaching provides the opportunity to merge the specialized knowledge and skills each teacher brings to the co-taught classroom, potentially benefiting all students in an inclusive environment (Villa et al., 2013).

Though the logic behind co-teaching is intuitively appealing, to date we know little about its effects. Only a handful of studies examining the efficacy of co-teaching include quantitative indicators of student outcomes (Murawski & Swanson, 2001; Cook et al., 2011), and most suffer from severe methodological limitations and/or are limited to only one or a few schools or classrooms.

We leverage longitudinal administrative data following the universe of public school stu-

dents and teachers in Massachusetts in order to provide the first causal estimate for the effect of co-teaching across a large public school system. For SWDs we find that attending a co-taught classroom leads to a significant but modest positive increase in a student's score on the state's English language arts (ELA) exam in elementary grades (three to five), but that attending a co-taught classroom in a secondary grade (six to eight, and 10) has a negative impact of similar magnitude. We also find evidence of a positive effect for SWDs in math that is primarily driven by students in elementary grades. Attending a co-taught classroom has a modest negative effect for students without disabilities in ELA. For both groups, the effect of co-teaching is independent of the impact from increasing the proportion of SWDs with whom they share classrooms.

Co-teaching has widespread support in much of the special education literature (e.g., Friend et al., 2010; Friend, 2015). The practice has also been widely recommended by state departments of education as a strategy for promoting inclusion of SWDs.<sup>3</sup> For example, the New York State Education Department has advocated that while co-teaching is not required, "school districts are strongly encouraged to phase this practice into its schools" (2008). Our findings from Massachusetts suggest that such broad endorsements of co-teaching are perhaps not warranted. Our results point to a more careful consideration of how co-teaching is deployed and suggest the need for additional research on the specific instructional co-teaching practices associated with positive student outcomes and how co-teaching is implemented at scale.

We also contribute to a wide body of research on policies intended to improve teacher quality throughout public school systems. The literature documenting the important role of teachers in producing student outcomes while in school and later in life has spawned a variety of educational reforms targeting instructional quality (see for instance Chetty et al., 2014). The teacher quality literature primarily focuses on the impact of individual teachers (Jones et al., 2019) and ways to improve their performance through additional training (for example, Kraft et al., 2018). Co-teaching is among a potential class of reforms, such as reducing class size, that alter the allocation of resources in a way that could better leverage current teachers' abilities.

The remainder of this paper proceeds as follows. Section 2 provides a brief introduc-

tion to co-teaching and an overview of existing scholarship on the topic. Section 3 describes the Massachusetts administrative data. Section 4 outlines our strategy for estimating the impact of co-teaching on student test scores. Section 5 reports our results. Section 6 presents a specification test showing that our primary results do not appear to be driven by dynamic selection into co-taught classrooms. Finally, Section 7 provides a brief summary of results and concludes.

## **II Co-Teaching**

Co-teaching allows schools to simultaneously meet multiple federal mandates related to educating SWDs. First, co-teaching allows schools to provide SWDs instruction in the least restrictive environment (LRE), as mandated by IDEA (formerly the Education for All Handicapped Children Act). For each SWD, the LRE is the setting in which the student is educated with students without disabilities to the maximum extent appropriate (Yell et al., 2011). Second, co-teaching allows schools to satisfy federal requirements associated with standards-based reforms. The No Child Left Behind Act (2002) included the requirement that SWDs receive instruction from highly qualified teachers in all academic content areas and make progress in the general education curriculum. Co-teaching provides SWDs access to highly qualified content area teachers and the general education curriculum, while simultaneously providing services from a special educator (Friend et al., 2010; Conderman, 2011).

Overall, the majority of studies on co-teaching describe the practice and offer recommendations for its implementation (e.g., Murawski & Dieker, 2004; Sileo, 2011) using information gleaned from observations of existing co-teaching programs (e.g., Dieker, 2001; Mastropieri et al., 2005) and interviews and surveys with practicing co-teachers (e.g., Austin, 2001; Keefe & Moore, 2004). As such, this literature is largely concentrated in two main areas: a) co-teaching program logistics<sup>4</sup> and b) co-teacher relationships.<sup>5</sup>

The intuitive appeal of co-teaching is widely recognized and reinforced by both general education and special education practitioners who work in co-taught classrooms. Overall, surveys

and interviews conducted with teachers indicate that both general education and special education teachers who participate in co-teaching have positive perceptions of co-teaching in general (Scruggs et al., 2007) and believe that it is related to positive outcomes for SWDs (Welch, 2000; McDuffie et al., 2009).

In particular, teachers have described greater academic benefits for students related to co-teaching over instruction in general education classes taught by a single general educator (e.g., Rice & Zigmond, 2000; Austin, 2001). Both general and special education teachers working in co-teaching situations attribute these academic benefits to the fact that students receive more individualized attention, teacher time, and assistance with the reduced student-teacher ratio in co-taught classes (Walther-Thomas, 1997; Trent, 1998; McDuffie et al., 2009). Teachers also report that all students benefit from the type of instruction offered in co-taught classes, which emphasizes remedial strategies and review (Austin, 2001) and includes support in organization skills and learning strategies (Trent, 1998).<sup>6</sup>

Despite this wide body of research describing co-teaching and its best practices, the evidence base regarding the effectiveness of co-teaching is scant. A recent observational study by Wexler et al. (2018) raised questions about the quality of instruction provided in co-taught classrooms. Drawing on over 2,000 minutes of observation in 16 co-taught classrooms, the authors found that little of the time in co-taught classrooms was devoted to instructional practices known to benefit students and in most cases the roles of special educators in co-taught classes were limited. Further, few studies of co-teaching have included measures of student achievement as outcomes and fewer still apply a method capable of producing causal estimates of treatment effects (Solis et al., 2012). There are a number of limitations on the conclusions that can be drawn from prior results due to study methodology. Only three prior studies were experimental and these took place within a single classroom or school (Murawski, 2006; Fontana, 2005); the rest either compared student outcomes across existing instructional models without random assignment or any sort of natural experiment (McDuffie et al., 2009; Hang & Rabren, 2009) or included no comparison group (Klinger et al., 1998).

In addition to the lack of evidence of co-teaching's effectiveness, the literature is lacking basic information about co-teaching that would be important in guiding policies surrounding its use. For example, existing research provides little data on the prevalence of co-teaching overall and its use in different subject areas or grade levels. Hock and Isenberg (2017) estimate that approximately 21% of students enrolled in tested grades in Washington DC's public schools receive instruction from more than one math teacher in a given year; 7% of these math teachers shared 100% of their students, suggesting they were co-teaching. We could identify no other published reports on the prevalence of co-teaching. Further, we could find no published information on the costs of co-teaching. It could be the case that co-teaching requires changes in overall staffing or introduces additional professional development expenditures to prepare teachers to work in co-taught classrooms.

We anticipate that the effectiveness of co-teaching might differ when applied in elementary or secondary classrooms. Because elementary teachers work in the same classroom with the same group of students each day, they are likely more prone to the kinds of deep, collaborative co-teaching relationships recommended in the special education literature. This is likely particularly true in reading instruction, where co-teaching could better foster the kinds of small group instructional interventions known to support SWDs (see the National Center for Intensive Intervention for a comprehensive list: <https://intensiveintervention.org/>).

We also explore the effects of co-teaching by student characteristics, including students' disability subcategories and whether there are differential effects across students with and without disabilities. Existing evidence would suggest that we cannot treat student disability categories interchangeably; these groups experience different learning trajectories and yearly teacher effects (e.g., Gilmour et al., 2019; Gilmour & Henry, 2018; Schulte & Stevens, 2015). Students from some "high-incidence disabilities" (e.g., specific learning disabilities, speech or language impairment, other health impairment) typically need less individualized support and are more likely to receive instruction in the general education classroom. Others, including students with autism and emotional behavioral disorders (EBDs), may need additional support for disruptive behavior

(Crosland & Dunlap, 2012; Harrison et al., 2012; Lane et al., 2006). Importantly, some evidence suggests that students without disability may exhibit lower academic performance when they have classmates with disabilities, particularly EBDs (Fletcher, 2010; Gottfried, 2014; Gottfried et al., 2016; Gottfried & Harven, 2015). By adding another teacher into the classroom, it is feasible that co-teaching arrangements could mitigate or perhaps even reverse any negative effect of inclusion on general education students. Unfortunately, we are aware of no prior research to date that has specifically evaluated the effect of co-teaching on the performance of students without disabilities.

### **III Data**

We utilize longitudinal administrative data for the universe of students and teachers from 2007-2008 through 2018-2019 made available by the Massachusetts Department of Elementary and Secondary Education (DESE). The analysis describing the prevalence and growth of co-teaching across the state includes data for all teachers in Grades K-12 each year. Analyses that evaluate the relationship between attending a co-taught classroom and later outcomes include data only on students enrolled in grades that were included in statewide standardized testing (Grades 3-8 and Grade 10).

The DESE data consists of four separate datasets: Student-level demographic and socioeconomic variables come from the Student Information Management System (SIMS) data, student test scores in math and ELA come from Massachusetts Comprehensive Assessment System (MCAS) data, teachers' job assignment classifications come from Education Personnel Management System (EPIMS) data, and information on student classroom assignments necessary to link students with their teachers comes from Student Course Schedule (SCS) data. For a detailed description of the data cleaning process, see the Appendix.

Using these data, we need to construct a student-year-level dataset with student test scores and indicators for the co-teaching status of the classrooms included. To accomplish the task, we first filter the SCS data to only include math and ELA courses and drop the duplicates in

terms of State Assigned Student Identifier (SASID), year, course, section, and term of instruction, which amounts to 0.05% of the sample. We then link these student-year-course observations to the EPIMS data by year, district, school, course, section, and term of instruction. We fail to match 2.2% of student-course-year observations to any educational staff in the EPIMS data. We also drop these observations. From there, we merge in student demographic characteristics from SIMS using the SASID, year, and school district. Approximately 7% of the sample of student-course-year observations do not have a corresponding student match in the SIMS data. The vast majority of these matching failures occur in 2011, which is the first year the state tracked student course work. We drop the unmatched observations. We then drop observations that correspond to courses taught in alternative education programs to facilitate merging in the MCAS data. We also restrict the sample to observations where the staff assigned to the course has a job assignment classified as either “teacher” or “co-teacher” in the EPIMS data.

One challenge in estimating the effects of co-teaching is that co-taught settings are often not clearly classified in administrative data. Although co-teacher is among teacher job assignment classifications, it is clear that administrators are not accurately identifying co-taught classrooms in all cases. For instance, as Table 1 shows, in only 11.5% of the classrooms where we observe two teachers is at least one teacher classified as co-teacher. In the remaining 88.5% of the ELA classrooms with two teachers, neither teacher is classified as a co-teacher. Table B4 in the Appendix shows that, within the ELA classrooms with at least one teacher classified as co-teacher, in only 40% of the observations we see both teachers classified as co-teachers, which is what we would expect if accurately classified, while in the remaining 60% of the observations we see one teacher classified as a co-teacher and the second teacher classified as a teacher.

In order to address this data issue, we classify co-taught classrooms according to two distinct definitions. Our primary classification, which we refer to from here as our “narrow” definition, identifies as co-taught any classroom for which we match two distinct teachers and at least one of the teachers is classified as a co-teacher. We argue that if an administrator views at least one teacher as a co-teacher then it is reasonable to suspect that he/she views the classroom as



co-taught.<sup>7</sup> We treat the estimated impact of co-teaching on student test scores according to this narrow definition as our primary measure of the effect of co-teaching.

We also separately identify classrooms for which we match two teachers and neither is classified as a co-teacher. We call this our “broad” definition for a co-taught classroom. Classifying these students into a separate category that we account for with a dummy variable in the regressions allows us to more directly compare the effect of attending a classroom that is co-taught according to the narrow definition to classrooms for which we match only a single teacher. We can also directly compare the estimates from the narrow and broad definitions, though we caution against interpreting the estimated effect of attending a co-taught classroom according to the broad definition as an estimate for the effect of co-teaching. If the two definitions equally identify co-taught classrooms then we should expect them to produce similar estimates for the impact of co-teaching on student performance. To the extent that the broader definition includes a mixture of co-taught classrooms and other classrooms that actually operate similarly to single-teacher classrooms though they are not classified that way, we would expect the estimate for the effect of co-teaching according to the broad definition to be biased toward zero. However, it is also possible that classrooms might have two assigned teachers for reasons other than co-teaching that we cannot fully anticipate, and thus we cannot predict the direction of potential bias. Thus, we focus on the narrow definition as our primary measure of co-teaching and we report the estimates from the broad definition for the sake of completeness.<sup>8</sup>

We emphasize that even in the best case we are not able to observe the practices of teachers working within co-taught classrooms. Prior research suggests wide variability in how teachers operate within co-taught settings (Murawski & Dieker, 2004; Sileo, 2011). Thus, our findings should be viewed as the estimated effect of co-teaching on average as implemented statewide in Massachusetts, which we argue has important implications for both policy and practice. Whether co-teaching is more or less effective when implemented in particular ways is an area for future research.

Figure 1 illustrates the growth in co-teaching that we identify according to both the nar-

row and broad definitions over time and geographically throughout Massachusetts. We observe a substantial increase in the proportion of students observed in co-taught classrooms throughout the state.

[FIGURE 1 ABOUT HERE]

Table 1 reports relevant descriptive statistics for the ELA sample.<sup>9</sup> The first three columns report mean characteristics for student-year observations in co-taught classrooms according to the narrow definition, according to the broad definition, and in single-teacher classrooms, respectively. The following three columns report the difference in means and the results of a *t-test* for inference comparing the two co-teaching classroom types, the broad co-teaching classrooms and the single-teacher classrooms, and the narrowly defined co-teaching classrooms and the single-teacher classrooms.

Though there are many statistically significant demographic differences in the characteristics of students in the narrowly and broadly defined co-taught classrooms, most are not of a meaningful magnitude. However, compared to those in co-taught classrooms according to the broad definition, students in the narrowly defined co-taught classrooms are substantially more likely to have an IEP and post substantially lower test scores.

The differences between the characteristics of students in the co-taught and single-teacher classrooms is more distinct. Compared to observations in single-teacher classrooms, observations of students in co-taught classrooms according to either definition are more likely to be African-American or Hispanic and to be eligible for free lunch. Students in co-taught classrooms also have substantially lower test scores and are much more likely to have an IEP than are students in single-teacher classroom settings.

[TABLE 1 ABOUT HERE]

## IV Empirical Method

Our goal is to estimate the causal effect of a student attending a co-taught classroom instead of a classroom headed by a single teacher on the student's academic performance as measured by standardized test scores. Our ideal strategy would compare the later test scores of students who were randomly assigned to attend a co-taught or single-teacher classroom. However, in practice schools adopted and implemented co-teaching nonrandomly. We suspect that a naive comparison of the outcomes of students who enrolled in co-taught and single-teacher classrooms would be biased by two sources of selection. First, there could be differences in the characteristics and leadership of schools that adopt co-teaching and those that do not. Second, schools might assign students to co-taught classrooms based on their academic ability or other attributes.

Our primary estimation strategy addresses these potential areas of selection by leveraging cross-student variation in the timing of assignment to a co-taught classroom within a regression model that holds constant student and school fixed effects. The model addresses the first source of selection bias by differencing out variation that is fixed across schools and addresses the second source of bias by differencing out fixed attributes for each student.

Our preferred regression model takes the form:

$$y_{ist} = \alpha + \sum_{k=1}^2 \beta_k \text{coteach}_{ist}^k + \sum_{k=1}^2 \lambda_k (\text{coteach}_{ist}^k \times \text{nosped}_{ist}) + \sum_{j=1}^J \psi_j X_{itj} + \delta_i + \phi_s + \gamma_t + \epsilon_{ist} \quad (1)$$

where  $y_{ist}$  is the standardized test score of student  $i$  while enrolled in school  $s$  during year  $t$ ;  $\delta_i$ ,  $\phi_s$ , and  $\gamma_t$  are student, school, and year fixed effects, respectively;  $X_{it}$  is a vector of time-variant observed characteristics for the student;  $\text{coteach}_{ist}^k$  is a dummy variable indicating whether the student attended a co-taught classroom where  $k = 1$  represents the narrow definition of co-teaching and  $k = 2$  represents the broader definition;  $\text{nosped}$  equals 1 if the student does not receive special education services and equals 0 if the student does receive such services;  $\epsilon_{ist}$  is a stochastic term

clustered by student; and  $\beta_k$ ,  $\lambda_k$ , and  $\psi_j$  are parameters to be estimated.

Note that we code the *nosped* to indicate that the student is not enrolled in special education so that when interpreting the result we highlight the effect of co-teaching on SWDs. The coefficients for  $\beta_1$  and  $\beta_2$  represent our estimate for the effect of a SWD attending a co-taught classroom according to the narrow or broad definition, respectively, relative to attending a classroom led by a single teacher. The estimate for the overall effect of a student without disabilities attending a co-taught classroom relative to attending a classroom led by a single teacher is found by the sum  $\beta_k + \lambda_k$ .

Causal interpretation of  $\beta_k$  and  $\lambda_k$  relies on the assumption that there are no unaccounted-for time-variant variables that are correlated both with the timing of a student's assignment to a co-taught classroom and their test score at the end of that year. That is, we assume that the timing of assignment to a co-taught classroom in a student's career is as good as random. Formally, the identifying assumption for Equation (1) is that:

$$E[y_{ist}|X_{it}, \delta_i, \phi_s, \gamma_t, coteach_{ist}] = E[y_{ist}|X_{it}, \delta_i, \phi_s, \gamma_t] \quad (2)$$

As we discussed in the Introduction, the adoption of co-teaching is largely motivated by the desire to include SWDs in general education classrooms along with students who do not have disabilities. The link between inclusion and co-teaching is apparent in the descriptive statistics reported in Table 1. From a policy perspective it is not obvious whether it is more relevant to isolate the effect of co-teaching from the effect of attending a more inclusive classroom setting. Thus, in addition to our primary estimate we report results from models that include a control for the percentage of students in the classroom who receive special education services, and also from models that interact the proportion of SWDs in the classroom with the variable indicating that the student does not have a disability. Because we are focused on estimating the effect of co-teaching itself, for our purposes we are primarily interested in whether and the extent to which controlling for this measure of inclusion impacts the estimated effect of enrolling in a co-taught classroom.

Nonetheless, the estimate for the relationship between the proportion of SWDs in a classroom and student outcomes has policy relevance as a limited measure of inclusion itself on student outcomes. However, we caution that convincingly estimating the causal effect of inclusive classroom settings itself requires a more in-depth analysis that treats with care issues of selection into more inclusive classroom settings beyond those that are within the scope of this current paper. Thus, though these estimates are intriguing, we do not claim them as the causal effect of inclusion and we leave that specific analysis for future research.

## A Specification Tests

We produce two analyses designed to evaluate the robustness of our findings and test underlying assumptions for giving a causal interpretation to the estimates. The first analysis alters the counterfactual condition used to estimate the effect of co-teaching for students with disabilities, and the second considers the possibility that our results are biased due to non-random sorting of students into co-taught classrooms based on time-variant characteristics.

The estimate for  $\beta_k$  derived from Equation (1) represents the average effect of attending a co-taught class for students with disabilities relative to the counterfactual single-teacher classroom that they would have attended otherwise. For students with disabilities, this counterfactual classroom might feasibly be either a general education classroom that they share with students without disabilities or a more isolated special education classroom, and so  $\beta_k$  represents the effect of attending a co-taught class compared to the weighted average of attending one of these other two class types. However, we might be concerned that for students with disabilities that effect of attending a co-taught class might differ by whether the alternative is attending a general or a specialized classroom. To address this issue, we estimate Equation (1) but add an indicator variable for whether the student is attending a specifically special education classroom. For these analyses, we limit the sample to include only students who are observed at some point attending a specifically special education classroom. Unfortunately, there is not a direct way to isolate special education classrooms in the data. Thus, we apply several potential definitions to identify special education

classrooms and consider the extent to which our estimate for  $\beta_k$  differs if we include an indicator variable identifying special education classrooms. We provide analyses that define a special education classroom as a class with only a single teacher (that is, not co-taught) and A) 80% or more students have an IEP or the teacher's job code is specific to teaching students with disabilities B) 80% or more students have an IEP or 25% of or more students with IEP and teacher's job code is specific to teaching students with disabilities C) 40% or more students have an IEP.

The most substantial threat to our identification strategy is the possibility that students are assigned to co-taught classrooms based on time-variant attributes in a way that violates the central assumption described in Equation (2).<sup>10</sup> For instance, administrators might assign a student to a co-taught classroom after an especially good or bad academic year.

Following the guidance from Angrist & Pischke (2008), we address the potential for such dynamic sorting into co-taught classrooms by comparing the results from Equation (1) with estimates from a regression that replaces the student fixed effect with a vector of observed time-invariant student characteristics ( $Z_i$ ) and the student's test score in the respective subject at the end of the prior year. Directly controlling for the student's test score in the previous year leads the model to estimate the effect of attending a co-taught classroom on student test score *gains*. Formally:

$$y_{ist} = \alpha + \sum_{k=1}^2 \beta_k \text{coteach}_{ist}^k + \sum_{k=1}^2 \lambda_k (\text{coteach}_{ist}^k \times \text{sped}_{ist}) + \sum_{j=1}^J \psi_j X_{itj} + \sum_{m=1}^M \delta_m Z_{im} + \chi y_{is(t-1)} + \phi_t + \gamma_s + \epsilon_{ist} \quad (3)$$

Causal interpretation of  $\beta_k$  and  $\lambda_k$  from estimating Equation (3) relies on the assumption that assignment to a co-taught classroom is essentially random conditional on the student's test score in the prior year.

$$E[y_{ist} | X_{it}, Z_i, y_{is(t-1)}, \phi_t, \gamma_s, \text{coteach}_{ist}] = E[y_{ist} | X_{it}, Z_i, y_{is(t-1)}, \phi_t, \gamma_s] \quad (4)$$

Since they rely on essentially opposing assumptions, the estimates from Equation (1) and Equation (3) produce a bound for the causal effect of enrolling in a co-taught classroom on student test scores (Angrist & Pischke, 2008). The trade-off in this approach is that Equation (3) now relies on  $y_{is(t-1)}$  and  $Z_i$  to sufficiently account for the influence of unchanging student factors that could be correlated with eventual assignment to a co-taught classroom.

In practice, the estimates for  $\beta_k$  and  $\lambda_k$  derived from Equation (1) and Equation (3) are not directly comparable because of sample differences and they employ variation from a different set of students to estimate the effect of co-teaching. First, the inclusion of the student's test score in the prior year eliminates from the sample all students who lack a test score in the prior year, including all third grade students. Second, because it includes a student fixed effect, the estimates of  $\beta_k$  and  $\lambda_k$  from Equation (1) are derived from students we observe in both a co-taught and non-co-taught classroom, which is not the case for the estimates from Equation (3). In order to assess the bounds of the the estimates, we thus compare the results from estimating Equation (3) to the results from estimating Equation (1) when the samples are restricted to include only students with a prior year test score and who are observed in both a co-taught and single-teacher setting during the sample period.

## V Results

Table 2 reports estimates for the impact of attending a co-taught classroom on student math and ELA scores on average for the full sample models. These regressions follow Equation (1) and include student, school, and year fixed effects. Columns 2 and 5 also add a control for the proportion of students in the classroom who have an IEP, and Columns 3 and 6 include an interaction between the percentage of students in the class with an IEP and an indicator for whether the student has an IEP.

[TABLE 2 ABOUT HERE]

The first row of the table reports estimates for the effect of attending a co-taught class-

room on the test scores of SWDs according to our narrow definition relative to attending a single-teacher classroom. On the ELA exam, we find a precisely estimated null effect of attending a co-taught classroom for students who have an IEP. We find a marginally significant positive effect from SWDs attending a co-taught classroom on the math exam. However, the magnitude of the effect is only 0.016 standard deviations, which is not economically meaningful.

Calculating the estimated effect of attending a co-taught classroom on students without disabilities according to the narrow definition requires adding the coefficients from the first and second rows. The *p-values* from *F-statistics* testing the null hypothesis that the sum of the coefficients is equal to zero are found at the bottom of the table. For students without disabilities, we find a precisely estimated null effect of attending a co-taught classroom on math outcomes. However, we find that attending a co-taught classroom according to the narrow definition leads to a statistically significant though modest (-0.035 standard deviation) decline in the ELA scores of students without disabilities.

The table also reports estimates for the effect of attending a co-taught classroom according to the broad definition. Recall that we believe that this category likely includes some co-taught classrooms and others that are not actually co-taught classrooms, and thus these estimates do not have a straightforward interpretation. On the math exam we find no significant impact of enrolling in a broad co-taught classroom, though the coefficients are precisely estimated. In ELA, for SWDs we find a more negative impact from attending a co-taught classroom according to the broad definition than from attending a co-taught classroom according to the narrow definition. We also find a statistically significant negative effect from attending a broadly defined co-taught classroom for students without disabilities, though the magnitude of the effect is insubstantial (0.006 standard deviations).

The latter columns of Table 2 show that accounting for the percentage of classmates with an IEP has little to no impact on the estimated effect of co-teaching on student test scores. Indeed, the coefficients on the narrow co-teaching variable and its interaction with IEP status are nearly unchanged when the model adds controls for the proportion of students in the classroom who



have a disability. This result suggests that the effect of co-teaching is independent of the effect of additional inclusion of SWDs in the general classroom setting.

Though it is not the central purpose of this paper, it is notable that taken at face value the results suggest that increasing the proportion of classmates who have a disability in a student's classroom is associated with a statistically significant decline in ELA scores for students without disabilities and a significant decline in math scores for both SWDs and for students without disabilities. However, the magnitudes of these effects are quite small. For example, the estimates in Columns 3 and 4 suggest that for a student without a disability a one standard deviation increase in the proportion of classmates who have a disability (an increase of about 0.10) is associated with a 0.002 standard deviation decrease in ELA scores and a 0.01 standard deviation decrease in math scores.

Figure 2 illustrates the coefficient and 95% confidence interval for the estimated effect of attending a co-taught classroom, according to our narrow definition, on student test scores overall and then separately for elementary and secondary grades.<sup>11</sup> The estimates for the impact of co-teaching on student test scores overall reported in Table 2 mask some meaningful heterogeneity in the effect by student grade level on the ELA exam. Indeed, the estimated effect of attending a co-taught classroom is nearly as positive for students in elementary grades as it is negative for students in secondary grades. Similarly, for students without disabilities we find a statistically and economically significant negative effect of attending a co-taught classroom in secondary grades, but no effect of enrolling in a co-taught classroom for students in elementary grades. We see less evidence of meaningful heterogeneity by grade level on the math exam.

[FIGURE 2 ABOUT HERE]

Figure 3 illustrates the results from models that separately measure the effect of attending a co-taught classroom according to our narrow definition for students with different disability classifications. The substantial reduction in the number of observations when the sample is restricted to students with particular classifications leads to less precise estimates, though the pattern of the results is informative.

Specific learning disability (SLD) is both the largest category within special education and the classification in which students are most likely to be assigned to a co-taught setting. Consistent with the estimated effect of co-teaching when we consider special education overall, on the ELA exam the effect for students classified as having an SLD is positive (0.063 standard deviations) in elementary grades but negative (-0.053 standard deviations) in secondary grades.

In most cases, the estimated impact of attending a co-taught classroom for students with disability classifications other than SLD are too imprecisely estimated to be informative. The results suggest that for students with an emotional disorder, attending a co-taught classroom could have positive effects in math, but, in secondary grades a substantial negative effect (about -0.161 standard deviations) on test scores in ELA. We also find some evidence of a positive effect of co-teaching on students with a health disorder in elementary grades. The results suggest that attending a co-taught classroom does not meaningfully impact the performance of students with other disability classifications.

[FIGURE 3 ABOUT HERE]

## **VI Results from Specification Tests**

Table 3 reports the results from our first specification test, in which we estimate models equivalent to Equation (1) except that they include an indicator for whether the student is attending a single-teacher class that is specific for students with disabilities and the sample is restricted to include only students who we observe at least once attending both a special education classroom. Recall that this analysis is meant to address the fact that students with disabilities who do not attend co-taught classes might either attend a general education class or a class specific to students with disabilities. Further, recall that because there is no clear way to identify special education classes we estimate models in which we identify such classes in one of three ways. The estimated impact of a student with disabilities attending a co-taught classroom found in Table 3 are quite similar to our main results reported in Table 2.<sup>12</sup> Thus, it does not appear that our main findings are significantly

impacted by the two potential counterfactual conditions for students with disabilities.

[TABLE 3 ABOUT HERE]

Figure 4 illustrates the results from our specification test comparing estimates from student fixed effects models with those from models that instead control for observed student demographics and a lagged dependent variable. Recall that the purpose of this test is to evaluate whether the estimated impact of co-teaching depends on whether we assume that students are sorted into co-caught classrooms based on fixed or time-variant characteristics. The two approaches make opposing assumptions about the potential for student sorting into co-taught classrooms, and thus the difference in their estimates provide a feasible range for the true treatment effect (Angrist & Pischke, 2008).

[FIGURE 4 ABOUT HERE]

The results of the specification test suggest that the estimates from the preferred student fixed effects model are not biased by dynamic sorting into co-taught classrooms. In the large majority of cases there is little observed difference in the coefficient estimate for the effect of co-teaching between the fixed effect and lagged dependent variable approaches, and in no case is the difference between the estimates statistically significant.

## **VII Summary and Conclusion**

We produce the first causal estimate of co-teaching on student academic achievement within a large public school system. Drawing on longitudinal administrative data from Massachusetts, we explore the impact of being assigned to a co-taught classroom in a given school year for both students with and without disabilities. The effect of co-teaching differs substantially in elementary and secondary grades. In elementary grades, co-teaching leads to modest benefits for SWDs, at least in ELA, and has no impact on the performance of students without disabilities. In secondary grades, attending a co-taught classroom leads to modest declines in student test scores, again especially in ELA, for

both SWDs and their classmates without a disability. The effect of co-teaching is independent of the effect of increasing the proportion of students in the classroom who receive special education services for both students with and without disabilities.

Our primary contribution is to add much needed empirical information about the impacts of co-teaching as implemented within a broad set of public schools. The vast majority of the existing studies on co-teaching are qualitative or descriptive in nature and draw on samples of a limited number of students and schools. This research has contributed to our understanding of the various forms that co-teaching can take, and the kinds of supports that would need to be in place to support positive co-teaching working relationships. However, the body of prior research lacks the evidence on the effects of co-teaching on student outcomes that could justify such a large number of states endorsing co-teaching as a service delivery model.

Our results point to the need for a more careful consideration of how co-teaching is taken up in schools. First and foremost, co-teaching may be more appropriate in early grades than in later grades. Our finding that co-teaching has negative effects for both students with and without disabilities in middle school aligns well with a recent observational study conducted by Wexler et al. (2018), where the authors found that students in middle school co-taught ELA classrooms received little exposure to instructional practices known to support SWDs' reading comprehension. Instead, the special educators in the study spent most of their time supporting the general education teacher. One possible reason for the positive effects in elementary school is that in the earlier grades instruction is organized in ways that facilitate productive co-teaching relationships. Unlike in middle school, where special educators are likely to support students across multiple classrooms, elementary teachers are more likely to work in a single classroom with a single co-teacher (Parker et al., 2017). In addition, it is common to see structured reading and mathematics programs in elementary school, which clearly delineate the roles and responsibilities of classroom educators. With more time and supports to implement co-teaching, elementary school teachers are likely better able to work in the ways envisioned by supporters of co-teaching.

Our analysis is limited to considering the immediate academic effects of co-teaching.

However, it is important to keep in mind that some of the most important justifications for moving SWDs into inclusive classroom settings are not academic but social. If co-teaching were to improve the social skills and friendships of SWDs then it might be worthwhile even in middle school where we find negative test score impacts. However, it is not clear why co-teaching itself would have positive effects on the socialization of SWDs independent of creating inclusive classroom settings. Further, that our findings for the academic impact of co-teaching differ from the common narrative serves as a reminder of the importance of research that rigorously measures treatment effects in order to confirm or challenge the conventional wisdom.

The primary limitation to our analysis is that we are only able to measure the average effect of co-teaching as currently implemented across a statewide school system without accounting for the specific relationships and practices used within the co-taught setting. Proponents of co-teaching might reasonably argue that our findings are muted by schools and teachers that have moved students into classrooms that have multiple teachers but the teachers do not apply the best practices necessary for co-teaching to be effective. From a policy perspective, however, our finding that co-teaching as it is currently implemented has at best modest positive academic effects restricted to elementary grades is highly relevant. Whether our results are driven by limitations that are inherent to co-teaching or the impact of co-teaching could improve by better application of best practices is an important area for future research.

To many, co-teaching is a solution to the long-present challenge of ensuring that a historically marginalized population of students receives access to general education alongside their peers. Given the widespread public acceptance of co-teaching, it is distressing that we know so little about the impact of co-teaching and of inclusion overall on the academic and social outcomes of both students with and without disabilities. Given the continued achievement gap between SWDs and students without disabilities (Chudowsky & Chudowsky, 2009; Newman et al., 2011; Schulte & Stevens, 2015; Schulte et al., 2016), it is critical that researchers provide policymakers with rigorous evidence to inform special education policy decisions. Our paper is one of a few recent studies that use large administrative data to measure the cause effects within the context of

special education (Schwartz et al. (2019); Ballis & Heath (2019); Setren (2019)). In finding that co-teaching may indeed be an appropriate solution for some students, we complement this emerging literature; and, we encourage the field to follow these examples in building a more robust evidence base surrounding special education.

## References

- Angrist, J. D. & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.
- Austin, V. L. (2001). Teachers' beliefs about co-teaching. *Remedial and Special Education, 22*(4), 245–255. <https://doi.org/10.1177/074193250102200408>
- Ballis, B. & Heath, K. (2019). The long-run impacts of special education. *EdWorkingPapers, 19*-151. <https://doi.org/10.26300/a6df-4266>
- Bessette, H. J. (2008). Using students' drawings to elicit general and special educators' perceptions of co-teaching. *Teaching and Teacher Education, 24*(5), 1376–1396. <https://doi.org/10.1016/j.tate.2007.06.007>
- Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*(9), 2633–79. <https://doi.org/10.3386/w19424>
- Chudowsky, N. & Chudowsky, V. (2009). *State test score trends through 2007-08, part 4: Has progress been made in raising achievement for students with disabilities?* Center on Education Policy. [https://www.cep-dc.org/cfcontent\\_file.cfm?Attachment=Chudowsky2Kober\\_STST07-08\\_P4.ProgressStudentswDisabilities.111609.pdf](https://www.cep-dc.org/cfcontent_file.cfm?Attachment=Chudowsky2Kober_STST07-08_P4.ProgressStudentswDisabilities.111609.pdf)
- Conderman, G. (2011). Middle school co-teaching: Effective practices and student reflections. *Middle School Journal, 42*(4), 24–31. <https://doi.org/10.1080/00940771.2011.11461771>
- Conderman, G. & Hedin, L. R. (2014). Co-teaching with strategy instruction. *Intervention in School and Clinic, 49*(3), 156–163. <https://doi.org/10.1177/1053451213496158>
- Cook, B. G., McDuffie-Landrum, K. A., Oshita, L., & Cook, S. C. (2011). Co-teaching for students with disabilities. In J. M. Kauffman, D. P. Hallahan, & P. C. Pullen (Eds.), *Handbook of Special Education* (pp. 187–208). New York, NY: Routledge.

- Cramer, E. & Nevin, A. (2006). A mixed methodology analysis of co-teacher assessments. *Teacher Education and Special Education*, 29(4), 261–274. <https://doi.org/10.1177/088840640602900406>
- Crosland, K. & Dunlap, G. (2012). Effective strategies for the inclusion of children with autism in general education classrooms. *Behavior Modification*, 36(3), 251–269. <https://doi.org/10.1177/0145445512442682>
- Dieker, L. A. (2001). What are the characteristics of “effective” middle and high school co-taught teams for students with disabilities? *Preventing School Failure: Alternative Education for Children and Youth*, 46(1), 14–23. <https://doi.org/10.1080/10459880109603339>
- Fletcher, J. (2010). Spillover effects of inclusion of classmates with emotional problems on test scores in early elementary school. *Journal of Policy Analysis and Management*, 29(1), 69–83. <http://dx.doi.org/10.2139/ssrn.1424191>
- Fontana, K. C. (2005). The effects of co-teaching on the achievement of eighth grade students with learning disabilities. *Journal of At-Risk Issues*, 11(2), 17–22.
- Friend, M. (2008). Co-teaching: A simple solution that isn’t simple after all. *Journal of Curriculum and Instruction*, 2(2), 9–19. <https://doi.org/10.3776/joci.2008.v2n2p9-19>
- Friend, M. (2015). Welcome to co-teaching 2.0. *Educational Leadership*, 73(4), 16–22.
- Friend, M., Cook, L., Hurley-Chamberlain, D., & Shamberger, C. (2010). Co-teaching: An illustration of the complexity of collaboration in special education. *Journal of Educational and Psychological Consultation*, 20(1), 9–27. <https://doi.org/10.1080/10474410903535380>
- Gately, S. E. & Gately, F. J., Jr. (2001). Understanding coteaching components. *TEACHING Exceptional Children*, 33(4), 40–47. <https://doi.org/10.1177/004005990103300406>
- Gilmour, A. F. (2019). Teacher certification area and the academic outcomes of students with



- learning disabilities or emotional/behavioral disorders. *The Journal of Special Education*, (pp. Advance online publication). <https://journals.sagepub.com/home/sed>
- Gilmour, A. F., Fuchs, D., & Wehby, J. H. (2019). Are students with disabilities accessing the curriculum? A meta-analysis of the reading achievement gap between students with and without disabilities. *Exceptional Children*, 85(3), 329–346. <https://doi.org/10.1177/0014402918795830>
- Gilmour, A. F. & Henry, G. T. (2018). Who are the classmates of students with disabilities in elementary mathematics classrooms? *Remedial and Special Education*, 41(1), 18–27. <https://doi.org/10.1177/0741932518789493>
- Gottfried, M. A. (2014). Classmates with disabilities and students' noncognitive outcomes. *Educational Evaluation and Policy Analysis*, 36(1), 20–43. <https://doi.org/10.3102/0162373713493130>
- Gottfried, M. A., Egalite, A., & Kirksey, J. J. (2016). Does the presence of a classmate with emotional/behavioral disabilities link to other students' absences in kindergarten? *Early Childhood Research Quarterly*, 36, 506–520. <https://doi.org/10.1016/j.ecresq.2016.02.002>
- Gottfried, M. A. & Harven, A. (2015). The effect of having classmates with emotional and behavioral disorders and the protective nature of peer gender. *The Journal of Educational Research*, 108(1), 45–61. <https://doi.org/10.1080/00220671.2013.836468>
- Hang, Q. & Rabren, K. (2009). An examination of co-teaching: Perspectives and efficacy indicators. *Remedial and Special Education*, 30(5), 259–268. <https://doi.org/10.1177/0741932508321018>
- Harrison, J. R., Vannest, K., Davis, J., & Reynolds, C. (2012). Common problem behaviors of children and adolescents in general education classrooms in the united states. *Journal of Emotional and Behavioral Disorders*, 20(1), 55–64. <https://doi.org/10.1177/1063426611421157>

- Jones, N. D., Bettini, E., & Brownell, M. (2019). Competing strands of educational reform policy: Can collaborative school reform and teacher evaluation reform be reconciled? *The Elementary School Journal*, 119(3), 468–486. <https://doi.org/10.1086/701706>
- Keefe, E. B. & Moore, V. (2004). The challenge of co-teaching in inclusive classrooms at the high school level: What the teachers told us. *American Secondary Education*, 32(3), 77–88.
- Klinger, J., Vaughn, S., Hughes, M. T., Schumm, J. S., & Elbaum, B. (1998). Outcomes for students with and without learning disabilities. *Learning Disabilities Research & Practice*, 13(3), 153–161.
- Kohler-Evans, P. A. (2006). Co-teaching: How to make this marriage work in front of the kids. *Education*, 127(2), 260–265.
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547–588.
- Lane, K. L., Carter, E. W., Pierson, M. R., & Glaeser, B. C. (2006). Academic, social, and behavioral characteristics of high school students with emotional disturbances or learning disabilities. *Journal of Emotional and Behavioral Disorders*, 14(2), 108–117. <https://doi.org/10.1177/10634266060140020101>
- Mastropieri, M. A., Scruggs, T. E., Graetz, J., Norland, J., Gardizi, W., & McDuffie, K. (2005). Case studies in co-teaching in the content areas: Successes, failures, and challenges. *Intervention in School and Clinic*, 40(5), 260–270. <https://doi.org/10.1177/10534512050400050201>
- McDuffie, K. A., Mastropieri, M. A., & Scruggs, T. E. (2009). Differential effects of peer tutoring in co-taught and non-co-taught classes: Results for content learning and student-teacher interactions. *Exceptional Children*, 75(4), 493–510. <https://doi.org/10.1177/001440290907500406>

- Murawski, W. W. (2005). Addressing diverse needs through co-teaching. *Kappa Delta Pi Record*, 41(2), 77–82. <https://doi.org/10.1080/00228958.2005.10532049>
- Murawski, W. W. (2006). Student outcomes in co-taught secondary english classes: How can we improve? *Reading & Writing Quarterly*, 22(3), 227–247. <https://doi.org/10.1080/10573560500455703>
- Murawski, W. W. & Dieker, L. A. (2004). Tips and strategies for co-teaching at the secondary level. *TEACHING Exceptional Children*, 36(5), 52–58. <https://doi.org/10.1177/004005990403600507>
- Murawski, W. W. & Lochner, W. W. (2011). Observing co-teaching: What to ask for, look for, and listen for. *Intervention in School and Clinic*, 46(3), 174–183. <https://doi.org/10.1177/1053451210378165>
- Murawski, W. W. & Swanson, H. L. (2001). A meta-analysis of co-teaching research: Where are the data? *Remedial and Special Education*, 22(5), 258–267. <https://doi.org/10.1177/074193250102200501>
- Newman, L., Wagner, M., Knokey, A.-M., Marder, C., Nagle, K., Shaver, D., & Wei, X. (2011). *The post-high school outcomes of young adults with disabilities up to 8 years after high school: A report from the national longitudinal transition study-2 (nlts2)*. National Center for Special Education Research. <https://ies.ed.gov/ncser/pubs/20113005/pdf/20113005.pdf>
- Nierengarten, G. (2013). Supporting co-teaching teams in high schools: Twenty research-based practices. *American Secondary Education*, 42(1), 73–83.
- Parker, A., Rakes, L., & Arndt, K. (2017). Departmentalized, self-contained, or somewhere in between: understanding elementary grade-level organizational decision-making. *The Educational Forum*, 81(3), 236–255. <https://doi.org/10.1080/00131725.2017.1314569>
- Ploessl, D. M., Rock, M. L., Schoenfeld, N., & Blanks, B. (2010). On the same page: Practical

- techniques to enhance co-teaching interactions. *Intervention in School and Clinic*, 45(3), 158–168. <https://doi.org/10.1177/1053451209349529>
- Pratt, S. (2014). Achieving symbiosis: Working through challenges found in co-teaching to achieve effective co-teaching relationships. *Teaching and Teacher Education*, 41, 1–12. <https://doi.org/10.1016/j.tate.2014.02.006>
- Rice, D. & Zigmond, N. (2000). Co-teaching in secondary schools: Teacher reports of developments in Australian and American classrooms. *Learning Disabilities Research & Practice*, 15(4), 190–197.
- Schulte, A. C. & Stevens, J. J. (2015). Once, sometimes, or always in special education: Mathematics growth and achievement gaps. *Exceptional Children*, 81(3), 370–387. <https://doi.org/10.1177/0014402914563695>
- Schulte, A. C., Stevens, J. J., Elliott, S. N., Tindal, G., & Nese, J. F. (2016). Achievement gaps for students with disabilities: Stable, widening, or narrowing on a state-wide reading comprehension test? *Journal of Educational Psychology*, 108(7), 925–942. <https://doi.org/10.1037/edu0000107>
- Schwartz, A. E., Hopkins, B. G., & Stiefel, L. (2019). The effects of special education on the academic performance of students with learning disabilities. *EdWorkingPapers*, 19-86. <http://www.edworkingpaper.com/ai19-86>
- Scruggs, T. E., Mastropieri, M. A., & McDuffie, K. A. (2007). Co-teaching in inclusive classrooms: A metasynthesis of qualitative research. *Exceptional children*, 73(4), 392–416. <https://doi.org/10.1177/001440290707300401>
- Setren, E. (2019). Targeted vs. general education investments: Evidence from special education and english language learners in boston charter schools. *Journal of Human Resources*, 0219–10040R2.

- Sileo, J. M. (2011). Co-teaching: Getting to know your partner. *Teaching Exceptional Children*, 43(5), 32–38.
- Solis, M., Vaughn, S., Swanson, E., & Mcculley, L. (2012). Collaborative models of instruction: The empirical foundations of inclusion and co-teaching. *Psychology in the Schools*, 49(5), 498–510. <https://doi.org/10.1002/pits.21606>
- Trent, S. C. (1998). False starts and other dilemmas of a secondary general education collaborative teacher: A case study. *Journal of Learning Disabilities*, 31(5), 503–513. <https://doi.org/10.1177/002221949803100508>
- Vaughn, S., Schumm, J. S., & Arguelles, M. E. (1997). The abcdes of co-teaching. *Teaching Exceptional Children*, 30(2), 4–10. <https://doi.org/10.1177/004005999703000201>
- Villa, R. A., Thousand, J. S., & Nevin, A. I. (2013). *A guide to co-teaching: New lessons and strategies to facilitate student learning*. Thousand Oaks, CA: Corwin Press.
- Walther-Thomas, C., Bryant, M., & Land, S. (1996). Planning for effective co-teaching: The key to successful inclusion. *Remedial and Special Education*, 17(4), 255–264. <https://doi.org/10.1177/074193259601700408>
- Walther-Thomas, C. S. (1997). Co-teaching experiences: The benefits and problems that teachers and principals report over time. *Journal of Learning Disabilities*, 30(4), 395–407. <https://doi.org/10.1177/002221949703000406>
- Welch, M. (2000). Descriptive analysis of team teaching in two elementary classrooms: A formative experimental approach. *Remedial and Special Education*, 21(6), 366–376. <https://doi.org/10.1177/074193250002100606>
- Wexler, J., Kearns, D. M., Lemons, C. J., Mitchell, M., Clancy, E., Davidson, K. A., Sinclair, A. C., & Wei, Y. (2018). Reading comprehension and co-teaching practices in

middle school english language arts classrooms. *Exceptional Children*, 84(4), 384–402.  
<https://doi.org/10.1177/0014402918771543>

Wilson, G. L. (2005). This doesn't look familiar! A supervisor's guide for observing co-teachers. *Intervention in School and Clinic*, 40(5), 271–275.  
<https://doi.org/10.1177/10534512050400050301>

Wilson, G. L. (2008). Be an active co-teacher. *Intervention in School and Clinic*, 43(4), 240–243.  
<https://doi.org/10.1177/1053451208314911>

Yell, M. L., Katsiyannis, A., & Bradley, M. R. (2011). The individuals with disabilities education act: The evolution of special education law. In J. M. Kauffman, D. P. Hallahan, & P. C. Pullen (Eds.), *Handbook of Special Education* (pp. 71–86). New York, NY: Routledge.

## VIII Figures and Tables

Figure 1: Change in the Share of Co-Taught Students by Massachusetts School Districts, ELA Sample

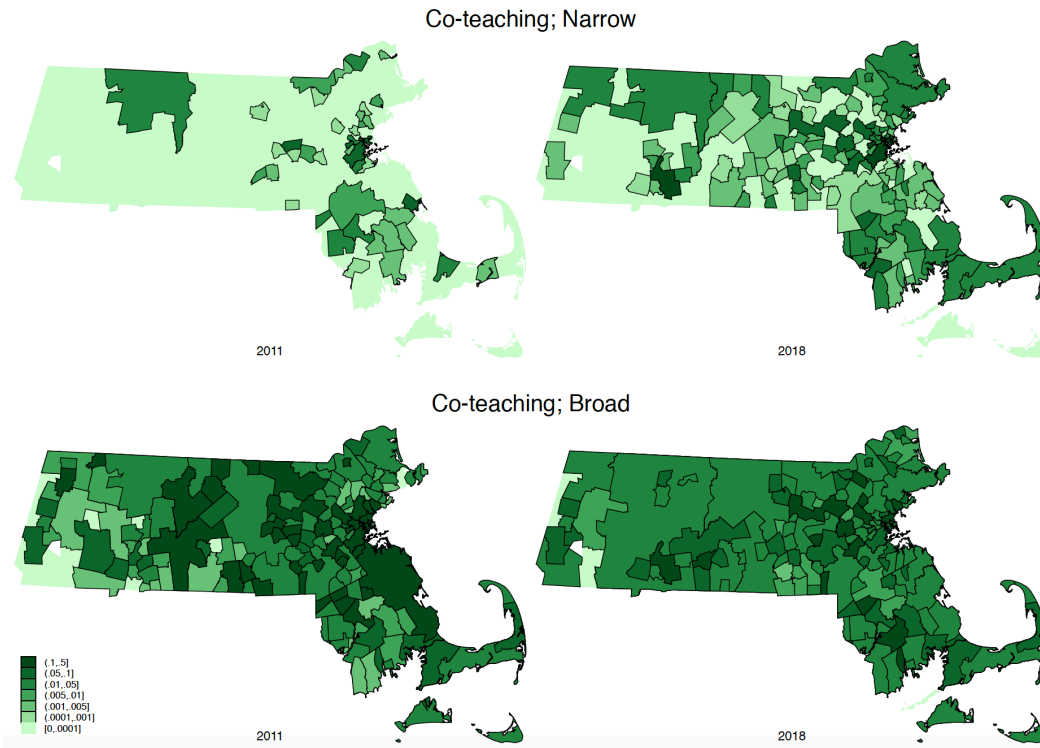


Figure 2: Effect of Co-Teaching (Narrow) by Grade Levels

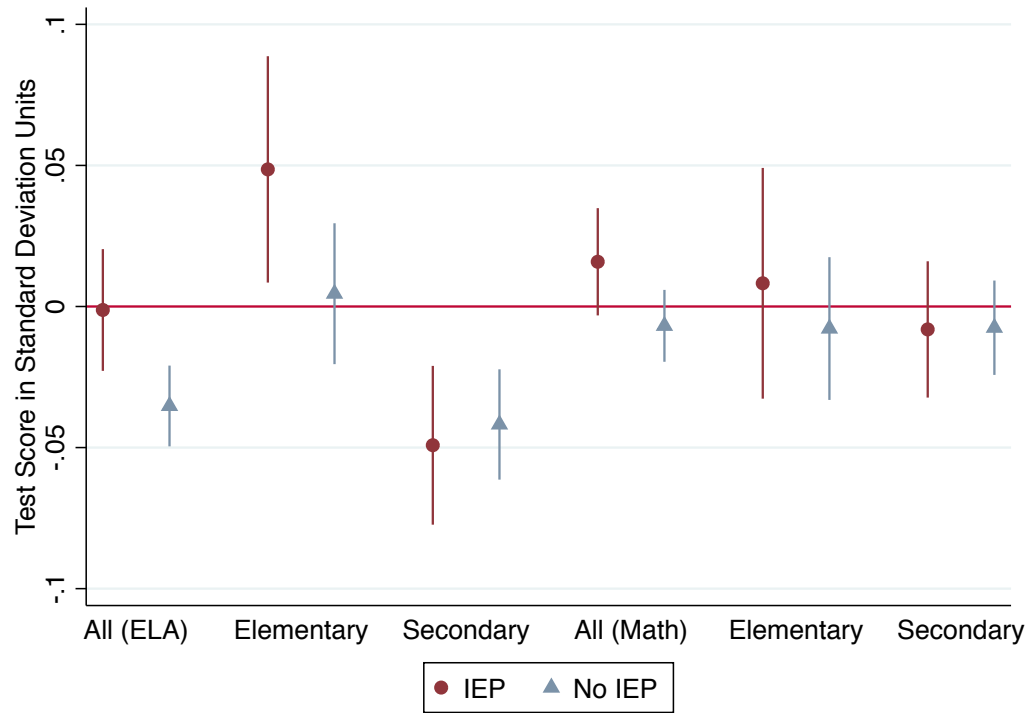


Figure 3: Effect of Co-Teaching (Narrow) by Disability Classifications

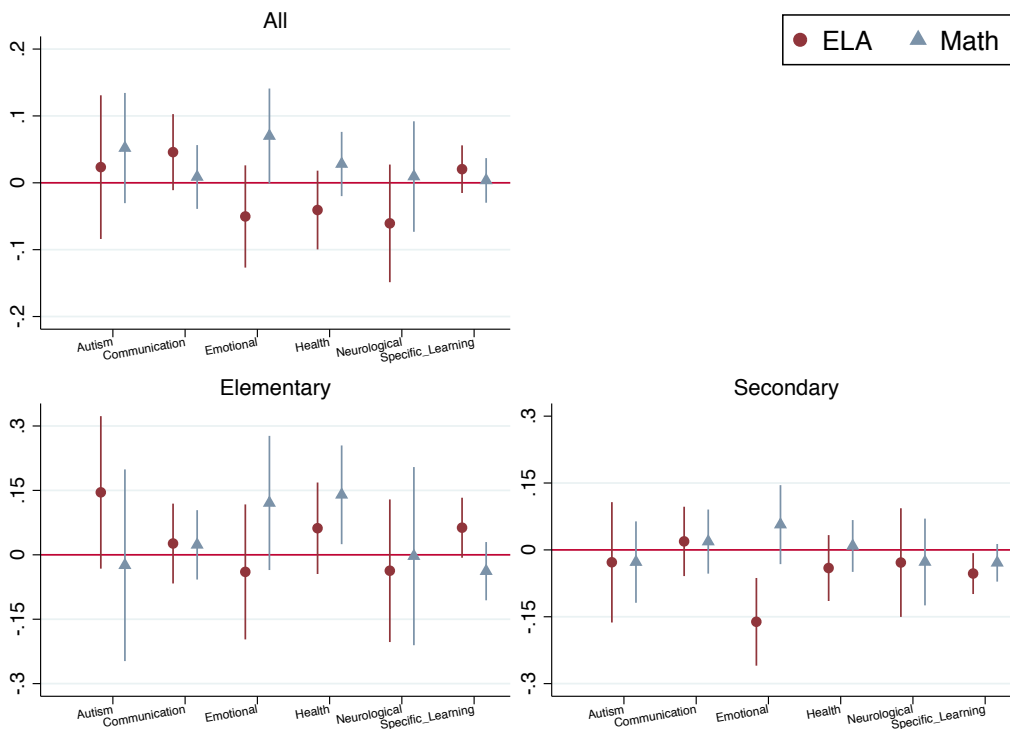




Figure 4: Effect of Co-Teaching (Narrow): Student Fixed Effect and Lagged Dependent Variable

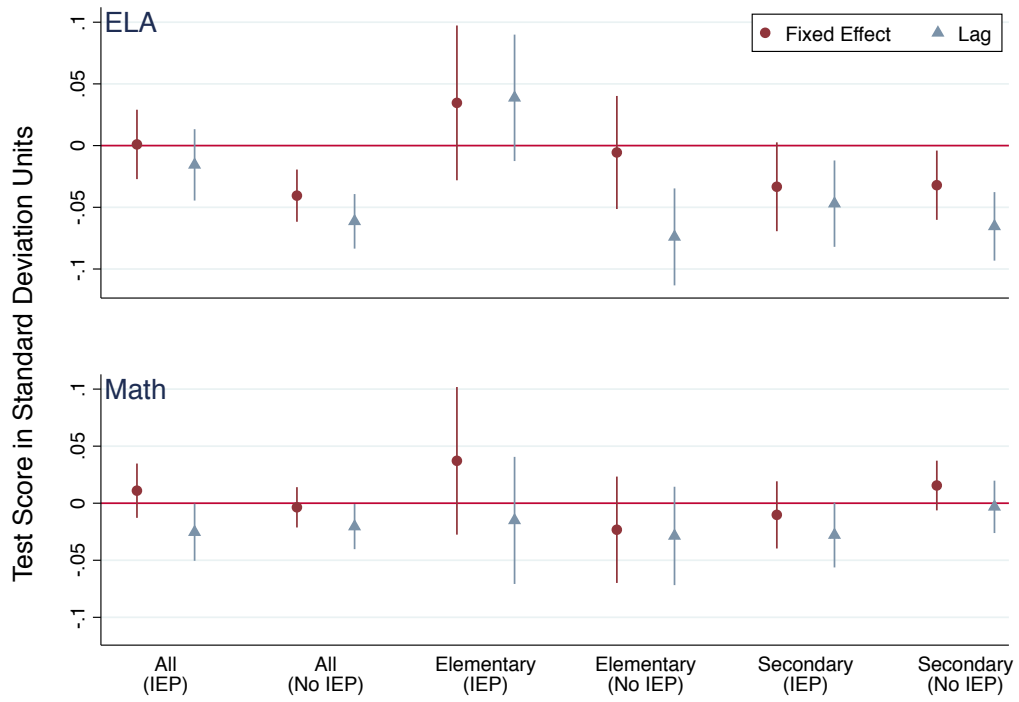


Table 1: Descriptives: ELA sample

|                   | (1)<br>Co-Teach; Narrow | (2)<br>Co-Teach; Broad | (3)<br>Single Teacher | (4)<br>Difference (1)-(2) | (5)<br>Difference (2)-(3) | (6)<br>Difference (1)-(3) |
|-------------------|-------------------------|------------------------|-----------------------|---------------------------|---------------------------|---------------------------|
| Female            | 0.451                   | 0.467                  | 0.496                 | -0.016***<br>(0.002)      | -0.029***<br>(0.001)      | -0.045***<br>(0.002)      |
| White             | 0.751                   | 0.772                  | 0.821                 | -0.021***<br>(0.002)      | -0.048***<br>(0.001)      | -0.069***<br>(0.001)      |
| Black             | 0.202                   | 0.180                  | 0.126                 | 0.022***<br>(0.001)       | 0.054***<br>(0.000)       | 0.076***<br>(0.001)       |
| Hispanic          | 0.271                   | 0.273                  | 0.162                 | -0.002<br>(0.002)         | 0.111***<br>(0.001)       | 0.109***<br>(0.001)       |
| Asian             | 0.071                   | 0.065                  | 0.077                 | 0.006***<br>(0.001)       | -0.012***<br>(0.000)      | -0.006***<br>(0.001)      |
| Free Lunch        | 0.524                   | 0.481                  | 0.312                 | 0.043***<br>(0.002)       | 0.169***<br>(0.001)       | 0.212***<br>(0.002)       |
| Reduced Lunch     | 0.046                   | 0.046                  | 0.039                 | 0.001<br>(0.001)          | 0.006***<br>(0.000)       | 0.007***<br>(0.001)       |
| IEP               | 0.387                   | 0.296                  | 0.193                 | 0.091***<br>(0.002)       | 0.103***<br>(0.001)       | 0.194***<br>(0.002)       |
| Autism            | 0.021                   | 0.014                  | 0.012                 | 0.006***<br>(0.000)       | 0.003***<br>(0.000)       | 0.009***<br>(0.000)       |
| Communication     | 0.051                   | 0.035                  | 0.024                 | 0.016***<br>(0.001)       | 0.011***<br>(0.000)       | 0.027***<br>(0.001)       |
| Emotional         | 0.028                   | 0.028                  | 0.014                 | 0.001<br>(0.001)          | 0.014***<br>(0.000)       | 0.015***<br>(0.001)       |
| Health            | 0.046                   | 0.035                  | 0.023                 | 0.011***<br>(0.001)       | 0.012***<br>(0.000)       | 0.023***<br>(0.001)       |
| Neurological      | 0.020                   | 0.014                  | 0.010                 | 0.006***<br>(0.000)       | 0.004***<br>(0.000)       | 0.010***<br>(0.000)       |
| Specific Learning | 0.104                   | 0.082                  | 0.052                 | 0.022***<br>(0.001)       | 0.030***<br>(0.000)       | 0.052***<br>(0.001)       |
| ELA Score         | -0.485                  | -0.375                 | 0.078                 | -0.110***<br>(0.005)      | -0.453***<br>(0.002)      | -0.563***<br>(0.005)      |
| Math Score        | -0.467                  | -0.352                 | 0.077                 | -0.115***<br>(0.005)      | -0.428***<br>(0.002)      | -0.544***<br>(0.004)      |
| Observations      | 92577                   | 711372                 | 4389954               | 803949                    | 5101326                   | 4482531                   |

Table 2: Effect of Co-Teaching on Students Test Scores

|                                 | ELA                  |                      |                      | Math                |                      |                      |
|---------------------------------|----------------------|----------------------|----------------------|---------------------|----------------------|----------------------|
|                                 | (1)                  | (2)                  | (3)                  | (4)                 | (5)                  | (6)                  |
| Co-Teaching; Narrow             | -0.001<br>(0.011)    | -0.001<br>(0.011)    | -0.001<br>(0.011)    | 0.016*<br>(0.010)   | 0.016<br>(0.010)     | 0.016*<br>(0.010)    |
| Co-Teaching; Narrow<br>× No IEP | -0.034***<br>(0.013) | -0.034***<br>(0.013) | -0.034***<br>(0.013) | -0.023**<br>(0.012) | -0.023*<br>(0.012)   | -0.023**<br>(0.012)  |
| Co-teaching; Broad              | -0.020***<br>(0.004) | -0.020***<br>(0.004) | -0.020***<br>(0.004) | -0.001<br>(0.004)   | -0.002<br>(0.004)    | -0.002<br>(0.004)    |
| Co-teaching; Broad<br>× No IEP  | 0.014***<br>(0.005)  | 0.014***<br>(0.005)  | 0.014***<br>(0.005)  | 0.003<br>(0.005)    | 0.004<br>(0.005)     | 0.004<br>(0.005)     |
| IEP Share                       |                      | -0.015**<br>(0.006)  | -0.008<br>(0.008)    |                     | -0.092***<br>(0.005) | -0.044***<br>(0.008) |
| IEP Share × No<br>IEP           |                      |                      | -0.021*<br>(0.012)   |                     |                      | -0.097***<br>(0.010) |
| Observations                    | 2590219              | 2590219              | 2590219              | 2575778             | 2575778              | 2575778              |
| r <sup>2</sup>                  | 0.835                | 0.835                | 0.835                | 0.865               | 0.865                | 0.865                |
| Student FE                      | Yes                  | Yes                  | Yes                  | Yes                 | Yes                  | Yes                  |
| School FE                       | Yes                  | Yes                  | Yes                  | Yes                 | Yes                  | Yes                  |
| Year FE                         | Yes                  | Yes                  | Yes                  | Yes                 | Yes                  | Yes                  |
| <i>F Test p values</i>          |                      |                      |                      |                     |                      |                      |
| Coteach; N + Interaction        | 0.000                | 0.000                | 0.000                | 0.274               | 0.291                | 0.294                |
| Coteach; B+ Interaction         | 0.017                | 0.016                | 0.014                | 0.378               | 0.415                | 0.465                |
| IEP Share + Interaction         |                      |                      | 0.001                |                     |                      | 0.000                |

Table 3: Effect of Co-Teaching with Special Education Classroom Controls, Restricted to Students Ever in a Special Education Classroom

|                                 | ELA                  |                      |                      | Math                 |                      |                      |
|---------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                                 | (1)                  | (2)                  | (3)                  | (4)                  | (5)                  | (6)                  |
| Coteaching; Narrow=1            | -0.004<br>(0.013)    | -0.003<br>(0.015)    | -0.016<br>(0.017)    | 0.013<br>(0.011)     | 0.008<br>(0.013)     | 0.001<br>(0.016)     |
| Co-Teaching; Narrow<br>× No IEP | -0.052***<br>(0.016) | -0.059***<br>(0.020) | -0.088***<br>(0.029) | -0.038***<br>(0.014) | -0.044***<br>(0.016) | -0.031<br>(0.027)    |
| Co-teaching; Broad              | -0.019***<br>(0.005) | -0.019***<br>(0.006) | -0.022***<br>(0.006) | -0.012**<br>(0.005)  | -0.012**<br>(0.005)  | -0.012**<br>(0.006)  |
| Co-teaching; Broad<br>× No IEP  | 0.018***<br>(0.006)  | 0.010<br>(0.007)     | -0.010<br>(0.010)    | 0.010*<br>(0.005)    | -0.005<br>(0.007)    | 0.008<br>(0.009)     |
| SPED Class                      | -0.004<br>(0.003)    | -0.010***<br>(0.003) | 0.004<br>(0.004)     | 0.002<br>(0.002)     | -0.002<br>(0.003)    | 0.009***<br>(0.003)  |
| SPED Class × No<br>IEP          | 0.006**<br>(0.003)   | 0.008**<br>(0.003)   | -0.000<br>(0.005)    | -0.003<br>(0.002)    | -0.011***<br>(0.003) | -0.014***<br>(0.004) |
| Observations                    | 1589319              | 826628               | 375846               | 1468548              | 846933               | 421559               |
| r <sup>2</sup>                  | 0.834                | 0.831                | 0.800                | 0.862                | 0.849                | 0.816                |
| Student FE                      | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  |
| School FE                       | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  |
| Year FE                         | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  | Yes                  |
| <i>F Test p-values</i>          |                      |                      |                      |                      |                      |                      |
| Coteach; N + Interaction        | 0.000                | 0.000                | 0.000                | 0.002                | 0.000                | 0.151                |
| Coteach; B + Interaction        | 0.770                | 0.037                | 0.000                | 0.468                | 0.000                | 0.507                |

**Note:** Each column represents models using different ways of identifying special education classrooms. In columns (1) and (4), a special education classroom is identified when the teacher has a job assignment of special education or when the share of IEP students is higher than 80%; In columns (2) and (5), a special education classroom is identified when the teacher has a job assignment of special education and the share of IEP students in the class is higher than 25% or when the share of IEP students is higher than 80%; In columns (3) and (6), a special education classroom is identified when the share of IEP students is higher than 40%.

## Notes

<sup>1</sup>National Center for Education Statistics, Digest of Education Statistics, 2017, Table 204.30

<sup>2</sup>National Center for Education Statistics, Digest of Education Statistics 2017, Table 204.60

<sup>3</sup>Examples of states promoting co-teaching include Texas (Texas Education Agency, n.d.), Oklahoma (Oklahoma State Department of Education, n.d.) Georgia (Georgia Department of Education, n.d.), and California (California Department of Education, 2019)

<sup>4</sup>The literature addressing logistical factors that influence the successful implementation of co-teaching has focused on the importance of establishing and maintaining a strong co-teacher relationship (e.g., Bessette, 2008; Gately & Gately, 2001; Keefe & Moore, 2004; Mastropieri et al., 2005; Ploessl et al., 2010). Often referred to as a professional marriage, the relationship between co-teachers is described as developing and evolving over time as teachers get to know one another and work together to solve problems and address issues as they arise (Gately & Gately, 2001; Sileo, 2011).

<sup>5</sup>Based on this work addressing co-teaching logistics and relationships, practitioner-focused journals are replete with “how-to” articles and books offering recommendations for developing co-teaching programs (Murawski, 2005; Murawski & Dieker, 2004), planning and implementing instruction in the co-taught classroom (e.g., Vaughn et al., 1997; Wilson, 2008; Conderman & Hedin, 2014) maintaining productive co-teaching relationships (e.g., Kohler-Evans, 2006; Sileo, 2011; Pratt, 2014), supervising co-teachers (e.g., Walther-Thomas et al., 1996; Wilson, 2005; Nierengarten, 2013), and evaluating co-teachers (Murawski & Lochner, 2011).

<sup>6</sup>General and special education co-teachers have also indicated that co-teaching positively contributes to their own professional development, helping them to improve their teaching practice as they share knowledge and skills and learn from their co-teacher (Walther-Thomas, 1997; Rice & Zigmond, 2000; Cramer & Nevin, 2006). For example, in semi-structured interviews conducted by Austin (2001), K-12 special education teachers indicated that co-teaching led to increased content knowledge and general education teachers indicated that it led to improved classroom management and curriculum adaptation skills.

<sup>7</sup>Results in Table B4 show that the characteristics of students enrolled in classrooms with either two or one teacher classified as a co-teacher are very similar, which provides us with additional confidence in the decision to merge these two definitions.

<sup>8</sup>We drop the student-year-course observations to which more than two teachers are assigned, which amounts to 0.5% of the sample.

<sup>9</sup>Table B1 in the Appendix shows similar results for the math sample.

<sup>10</sup>See the Appendix tables for additional robustness tests. In particular, we present results from models that exclude the school fixed effect, add controls for a particular course (e.g. Algebra I), or add controls for whether the classroom included a paraprofessional or other classroom aid. The main results are robust to each of these specifications.

<sup>11</sup>See Table B5 and Table B6 in the Appendix for the relevant coefficients and standard errors illustrated on the figure.

<sup>12</sup>Appendix Table B11 in the Appendix reports results from estimating Equation (1) on the select sample of students who we observe at least once in a special education classroom. The results are very similar to those reported in 2.