The Spillovers of Child Disability on Peers' Education

Sofía Sierra Vásquez[1], Alice Dominici[3,2], Nicoletta Balbo[2], and Massimo Anelli[2]

[1]University of St. Gallen

[2]Bocconi University

[3]IMT School for Advanced Studies Lucca

October 10, 2025

Preliminary and incomplete. Please do not circulate.

Abstract

We study the causal effect of having classmates with disabilities on schooling outcomes. The use of Chilean administrative data covering more than 1 million students allows us to overcome selection issues in existing studies and to conduct a causal analysis despite the low incidence of disabilities in the population. We find that exposure to disability decreases primary school test scores, with mathematics grades experiencing more lasting effects. We document strong heterogeneity in effects by type of disability: negative effects are driven by intellectual disabilities, whereas language disorders cause improvement in Spanish scores, suggesting positive spillovers from targeted support measures. Moreover, negative spillovers are strongly or fully moderated by maintaining the same teacher for more than one grade, a relatively simple intervention recommended for a wide array of other policy targets.

1 Introduction

As of 2025, 186 countries have ratified the United Nations Convention on the Rights of Persons with Disabilities (2006), aimed at promoting inclusion in every aspect of daily life. As a consequence, countries all over the world have implemented – or are planning to implement – educational reforms that integrate students with disabilities into mainstream schools, while providing them with the extra support needed to learn to the best of their capacity. Implementing such reforms is no easy task: it requires training specialized support teachers, and preparing teachers in mainstream schools. The transition to inclusive education must be designed to ensure that the contact between students with and without disabilities does not translate into additional barriers to learning for either group, but rather constitutes a source of social and personal enrichment. How do students without disabilities perform in school after exposure to disability? Despite the high policy relevance, the available literature on disabilities from pedagogy focuses on small convenience samples in very specific institutional settings.[1]

In this paper, we address this question by considering the case of Chile, which implemented a series of reforms for inclusive education from 2009 to 2015. By combining administrative data on the universe of students entering primary education during this period, we estimate the causal effect of exposure to disability on peers' test scores in Spanish and Mathematics (the two main school subjects). We rely on two identification strategies that return the same results, following a common approach in the vast education economics literature on peer effects (among the most recent contributions, see Carrell et al., 2018; Anelli et al., 2023). The first strategy relies on idiosyncratic variation in the prevalence of disabilities across cohorts. The intuition is the following: for a given school, grade, and fixed set of teachers, students who start primary school in two subsequent years will be selected into different exposure to disability simply because the number of children with disabilities varies across cohorts. The second strategy conditions on school and cohort, and compares students

[1] For example, see the reviews by Nowicki and Sandieson (2002) and Kalambouka (2007): they show that most studies from the pedagogy and education literature rely on small samples and provide qualitative or associational results.

assigned to different class letters (which correspond to a fixed set of teachers and classmates): since this assignment happens in first grade, before student characteristics become apparent, there is no evidence that students without disabilities are selected into exposure or different class environments based on them. Under both strategies, we find balanced individual and class-level covariates between exposed and non-exposed students.

Our main results, based on standardized test scores of a subset of Chilean students enrolled in 2nd to 8th grade between 2012 and 2019 ($N \approx 2,000,000$), indicate that exposure to disability decreases standardized Math and Spanish tests' scores by approximately 2%-3% of a standard deviation. The effects tend to disappear over time for Spanish and persist for Math, highlighting the different role played by cumulative learning in these subjects. The magnitudes of these effects are similar to findings on exposure to learning deficits (e.g., dyslexia, ADHD) in a study focused on Switzerland (Balestra et al., 2022). A heterogeneity analysis reveals that negative spillovers are mostly driven by children with intellectual disabilities, who also constitute the vast majority of students with disabilities. However, teacher looping, already known to positively affect students' cognitive outcomes (Hill and Jones, 2018, e.g.,), is able to moderate negative spillovers and even revert their sign in the case of Math, constituting an important policy implication not limited to the case of inclusive education.

The paper proceeds as follows. Section 2 details our contribution and hypotheses, Section 3 summarizes relevant details about the institutional background and the reforms on inclusive childcare ocurring during the studied period, Section 4 presents the data, Section 5 introduces our regression models and causal identification strategy, Section 6 presents our preliminary results. Finally, Section 7 concludes and discusses future extensions.

2 Contributions and hypothesis

While a United Nations convention pushes the majority of world countries to switch from segregated to inclusive education, there are no large, causal studies on the effects this has on peers without disabilities. To our knowledge, there are only two large empirical studies available. Ruijs et al. (2010), focusing on the Netherlands, is an associational study: as such, it does not identify causal effects that can directly inform policies. For instance, associational studies cannot distinguish between the effect of exposure to disabilities and selection factors, such as children with disabilities being systematically included in smaller and better-performing classes. Balestra et al. (2022), conducted in Switzerland, defines special needs students as those who have learning difficulties but can attend mainstream education with extra support (e.g., dyslexia, dyscalculia, ADHD). While providing causal results, Balestra et al. (2022)'s study is not informative for more severe physical and intellectual disabilities, which are the focus of the 2006 UN Convention and, thus, all subsequent educational reforms. Our main contribution will be to provide causal evidence on the inclusion of students with severe disabilities.

Our preliminary results indicate that exposure to disability has a negative effect on peers' schooling outcomes. Several theories can predict a negative effect, especially for disabilities that can cause disruption to class activities, such as certain types of cognitive impairments. Indeed our data, obtained by the Chilean Ministry of Education, confirms that the majority of children with permanent disabilities have an intellectual disability. Negative associations have also been found in qualitative studies reviewed by Kalambouka (2007), as well as causal studies in economics focused on different peer effects that cause disruptions, such as Carrell et al. (2018) and Carrell and Hoekstra (2010), who study exposure to children who experience trauma at home.

Indeed, we do not exclude that a positive effect might arise from our heterogeneity analyses. A positive effect might be driven by increased teachers' efforts to monitor the environment in selected schools; while in general a positive association could be driven by a selection mechanism by which students with disabilities are assigned to better teachers, smaller classes or one extra teacher (dedicated to special needs students), our identification strategy will account for teacher fixed effects. This means that we can test whether such selection is in place, but we can also exclude that they drive the estimated causal effects of exposure to disability. Survey data at the teacher level will also inform our analysis of mechanisms in this regard: they are compiled on the date of national standardised tests used for our grade outcomes, and contain questions on episodes of bullying and discrimination, students' feelings about attending school, teachers' pedagogic approach and impressions of the class environment.

3 Institutional background

In Chile, the school system is organized in two compulsory cycles: primary school, which lasts 8 years, and secondary school, which lasts 4 years. These can be provided by three types of schools: fully public, fully private, or voucher schools, which are publicly funded but privately managed schools. During our study period, 36% of first graders were enrolled in public schools, 56% in voucher schools, and 8% in private schools.

Admission to primary school has historically been a decentralized process in which schools are allowed to set their own admission criteria. These are typically based on students' potential achievement and socioeconomic background (Carrasco et al., 2017). For this reason, our analysis focuses only on within-school variation (either across cohorts or across classrooms) in exposure to classmates with disabilities. About 60% of first graders enroll in schools with more than one classroom per grade. In these cases, classroom assignment within a school has also been a decentralized process, where the allocation of children to classrooms seems to be as-good-as-random in about 60% of schools (Valenzuela Barros and Toledo Román, 2012).

During their school years, children regularly take a nationwide, standardized, low-stakes exam called SIMCE. This is a multiple-choice test on students' knowledge of math and Spanish. Every year since 1988, all Chilean students in a selected grade must take the SIMCE at the end of the school year. This implies that most cohorts of Chilean students took the test at least once. The test was temporarily suspended in 2020 and 2021 due to the COVID-19 pandemic.

3.1 Reforms for inclusive education

The history of inclusive education in Chile dates back to the 1990s, when a series of Decretes[2] established norms for the implementation of inclusive education programs by creating subsidies and training programs for schools receiving students with intellectual, sensory, motor disabilities, and speech disorders. As a result, in the early 2000s, about 8% of primary schools had students with special needs; nevertheless, they were usually taught in segregated groups (Godoy et al., 2004).

In the mid-2000s, in line with the UN Convention on the Rights of Persons with Disabilities (2006), the Chilean government stepped up its efforts to promote more inclusive education. As a result, the Inclusive Education Program ("Programa de Inclusión Escolar"), in short PIE, was created.[3] The program differentiated between children with *transitory* special needs, such as ADHD or deslexia, and *permanent* special needs. The program encouraged mainstream schools to participate in the program by establishing guidelines for diagnostic assessments, resource allocation, and the promotion of collaborative and inclusive classroom practices.

While PIE incentivized the integration of children with disabilities into mainstream schools, the funding structure and regulation of the general school system limited the effectiveness of PIE initiatives, as schools were still allowed to select students based on their potential academic performance and socio-demographic characteristics. This changed in 2015 with the creation of the *School Inclusion Law* (Law 20.845). The law prohibited discrimination against students in publicly funded schools, and it created a centralized admissions platform that matches parents' and students' preferences with available places in schools.

4 Data

Exposure to disability To perform our analysis we use administrative data on the universe of Chilean primary school students. Pseudonymized data on enrollment in any type of

[2] Decrete 490/90, Decrete 01/98, Decrete 291/99, and Decrete 374/99

[3] PIE is regulated by Law 20.201 (2007) and Supreme Decrete 170/2009.

school (public, voucher, or private) is publicly available for all years since 2004. We complement these data with restricted access information provided by the Ministry of Education on the *Special Needs* status of students. This additional information allows us to define students with a disability as those that require *permanent* support in education (according to the 2015 *School Inclusion Law*). We exclude from this definition students who require temporary aid, who typically have learning difficulties that can be overcome after a first period of special support courses. Our restricted access data is based on diagnoses at the individual level, but only reported macro-categories of disabilities for privacy reason. We can distinguish between children that present severe speech disorders, autism, sensory, motor, and intellectual disabilities, and have information on the co-occurrence of multiple disabilities.

We then measure *exposure to disability* with a binary indicator that takes value one if a student has a non-repeater classmate with a permanent disability when they first enter the first grade of primary school. As shown in Figure 1, students with disabilities are more likely to repeat grades compared to their peers, especially in the first years of school. This makes peers' exposure to disability endogenous to teacher characteristics in first grade and to potentially unobserved school and student characteristics in higher grades. For this reason, our exposure measure focuses on exposure to non-repeaters in their first grade of primary school.

Finally, note that unlike other studies on peer effects that use share measures – e.g., focusing on exposure to girls, immigrants, children with a violent family background – the occurrence of child disability is much rarer, implying that the majority of classes are not exposed and, when they are, the share of students with disabilities remains low. In econometric terms, child disability resembles the case of count data and, therefore, using shares would not have a clear interpretation. Figure 2 illustrates the distributions of the absolute number and the share of students with disabilities.

Outcome variables and moderators SIMCE test scores were provided by the Chilean governmental agency for the quality of education (*"Agencia de Calidad de la Educación"*). Table 1 summarizes the grades and cohorts tested in each year between 2007 and 2019.

During the test days, survey data on classroom environment is also collected both from students and teachers. This survey data will inform future analyses of mechanisms related to episodes of bullying, discrimination, and disruption, students' feelings about attending school, teachers' pedagogic approach, and impressions of the **class environment**.

Crucially, rich administrative data allows us to match students to their **teachers and their characteristics** and overall years of experience. We created a panel that follows teachers from 2003 to 2019. This allows to compute each teacher's prior experience with any child with disability. Since this is only based on the years of data we observe, it should be seen as a lower bound of the actual prior experince teachers may have. Additionally, we compute the number of years that the same teacher has taught a particular child with disability. All this information allows us to understand the moderating role of teachers' learning process in the integration of children with disabilities.

Finally, we will use restricted access school-level data on the use of subsidies on disability-specific **expenditures**. This data allows us to identify the resources used by schools each year on salaries for specialized special education staff, as well as on physical infrastructure and training for the entire school community on disability-related issues, that we will employ to perform heterogeneity analyses based on the intensity and type on spending by school.

Sample restrictions For our analyses, we focus on cohorts 2009 to 2015. The rationale for building our sample with these cohorts is two-fold. On the one hand, we want observe them taking the SIMCE at least once. This restrict our sample for the recent years, since cohorts 2016 to 2018 have never taken the test as a result of the COVID-19 suspension of the test. On the other hand, we need to be able to identify children who at some point of their school life will be considered to have a permanent disability. Since the data on the special needs is only available since 2015, this restricts our data in the earlier years, as students who started primary school in 2008 would be in the last year of elementary school in 2015, leaving us a very narrow timeframe to identify them. Figure 3 presents the evolution of our measure of exposure over the years. As the figure shows, there is a steady increase in the percentage of students exposed to classmates with disabilities between cohorts 2005 and 2009. Since we cannot be sure whether this is due to mismeasurement or to an actual increase in exposure, we exclude those years from the main analyses to avoid biasing our estimates towards zero. Nevertheless, we perform robustness checks using all the available cohorts. Given our main sample, Figure 4 shows that, conditional on having a disability, about 80% of children in our sample have intellectual disabilities and that the share of children diagnosed with autism increases steadily – a pattern confirmed anecdotally also by the Ministry of Education.

Figure 1: Probability of repeating a grade by disability status: primary school

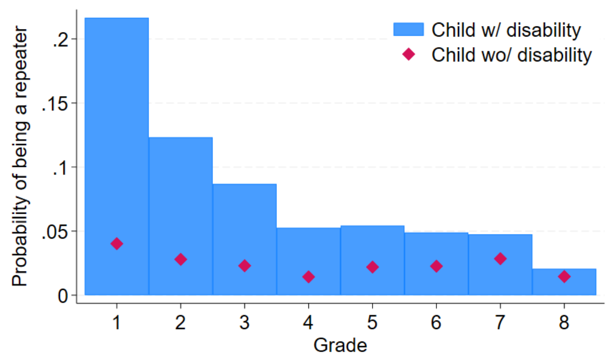
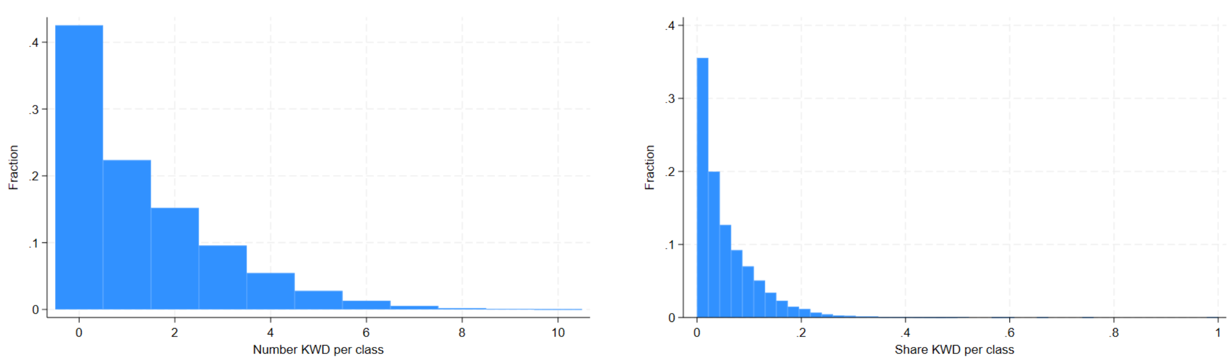


Figure 2: Exposure to kids with disability (KWD) per class: distribution



(A) Absolute number

(B) Share (0-1)

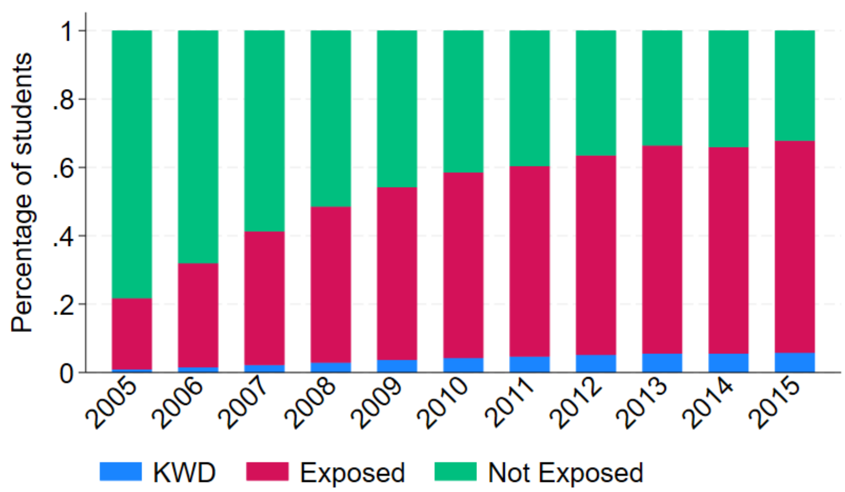
Notes: The figure displays the distribution of the absolute number of students with disabilities (Panel A) and the share of students with disabilities (Panel B) in first grade.

Table 1: Tested grades and cohorts by year - SIMCE

	Cohort											
Year	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
2007	4											
2008		4										
2009			4									
2010				4								
2011	8				4							
2012						**4**		**2**				
2013	10		8		6		**4**		**2**			
2014		10		8		**6**		**4**		**2**		
2015			10		8		**6**		**4**		**2**	
2016				10				**6**		**4**		
2017					10		**8**				**4**	
2018						10				**6**		**4**
2019									**8**			

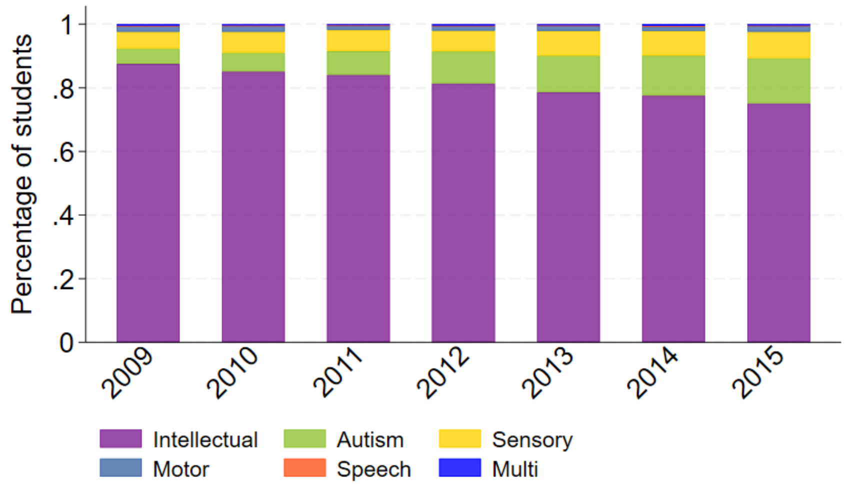
Note: Cohorts are identified by the year in which students start 1^{st} grade of elementary school. The table shows in which grades different cohorts took the SIMCE. Grades highlighted in bold font are the ones used for our analysis. Finally, note that second graders are only tested on their Spanish knowledge.

Figure 3: Probability of exposure to a classmate with disability in first grade



Notes: The figure displays the evolution of exposure to classmates with disabilities from 2005 to 2015. Our studied period corresponds to 2009-2015.

Figure 4: Percentage of students by type of disability



Notes: The figure shows how the types of disabilities that students were exposed to changed during the studied period.

5 Empirical strategy

We estimate the following regressions on a panel of students without disabilities (peers):

$$Score_{i,l,s,t} = \phi_{l,s} + \eta_t + \beta \text{Exp\_Dis}_{l,s,t} + \boldsymbol{\delta_1 X}_{i,l,s,t} + \boldsymbol{\delta_2 Z}_{l,s,t} + \boldsymbol{\delta_3 \Omega}_{s,t} + \varepsilon_{i,l,s,t} \tag{1}$$

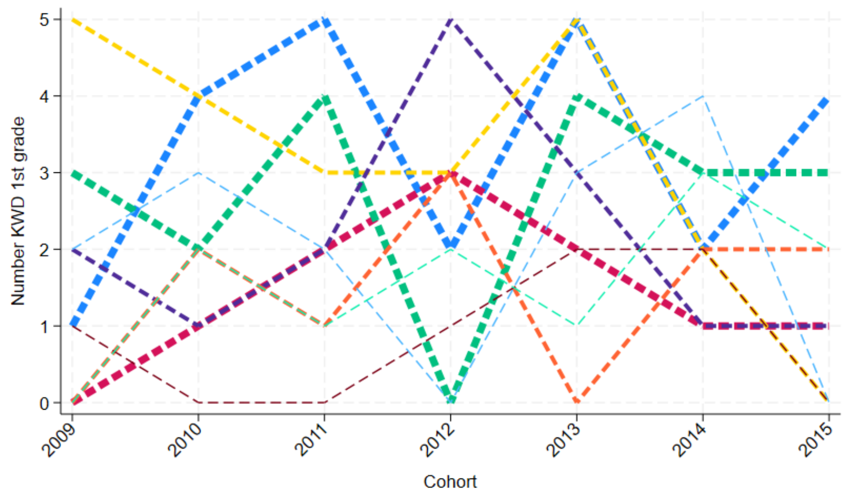
$$Score_{i,l,s,t} = \psi_{s,t} + \beta \text{Exp\_Dis}_{l,s,t} + \boldsymbol{\delta_1 X}_{i,l,s,t} + \boldsymbol{\delta_2 Z}_{l,s,t} + \boldsymbol{\delta_3 \Omega}_{s,t} + \varepsilon_{i,l,s,t} \tag{2}$$

Where Equation (1) corresponds to the cross-cohort identification strategy and Equation (2) to the cross-classroom one. The outcomes include Math and Spanish SIMCE scores. The two subjects are considered separately. We perform separate analyses for each observed grade, and add one regression for all grades pooled together, controlling for grade fixed-effects. Subscript i denotes individual students without a disability (peers), subscript s denotes schools, subscript t denotes first-grade cohorts (that have a 1-to-1 correspondence with years in this setting), and subscript l stands for "letter", i.e. it denotes a class group that, in most cases, has fixed class-group composition and fixed teachers for the entire education cycle. The treatment variable $\text{Exp\_Dis}_{l,s,t}$ varies by class-group and is a dummy equal to 1 if there is a classmate with a permanent disability in the first year and 0 otherwise. Note that since exposure is not time-varying after conditioning on each equation's fixed effects, we cannot control for student fixed effects. Instead, vectors X, Z, and Ω include time-varying individual, class, and school-level covariates, respectively. X includes a dummy for being a repeater, dummies for both parents' vocational or college education, and dummies for the family's income quintile. Z includes class size and class-level shares of students in each income quintile. Finally, Ω adds school size (number of students) and school-level shares of students whose parents have vocational and college education, separately for mothers and fathers.

Causal identification relies on the inclusion of fixed effects $\psi_{l,s}$ and η_t in Equation (1), and $\phi_{s,t}$ for Equation (2), which capture respectively school-by-cohort fixed effects and teachers-group fixed effects (i.e., identifying the class group that sticks together over the entire cycle of education and capturing teachers assigned to that group that do not change). In Equation (1), the variability that the model exploits, namely the residual variability after the inclusion of fixed effects, is the change in exposure to disability across cohorts. The thought experiment is the following: within the same grade (e.g. first year of high school), in a given school, and class-group (which includes a fixed set of teachers), the number of students with disabilities will differ between the cohorts that start in 2010 and 2011 simply due to idiosyncratic differences in the prevalence of disabilities, because the model specification accounts for all endogeneous selection.

Figure 5 displays the cross-cohort variation in the number of students with disabilities in first grade for nine randomly selected school-letter groups, to illustrate the absence of clear self-selection patterns. Table 2 shows that including school-by-cohort fixed effects and teachers-group fixed effects captures most of the selection based on observables. Note that while exposed children are significantly less likely to have parents with more than high school education or to belong to higher income quintiles, the size of these differences is in all cases less than 1 percentage point and are, therefore, not *economically* significant. We attribute the statistical significance of these differences to the large population size used in

Figure 5: Random cross-cohort exposure within school-letter groups



Notes: The figure displays the number of students with disabilities in first grade for nine randomly selected school-letter groups – the groups used to attain causal identification in our cross-cohort analysis, corresponding to Equation (1).

our analyses. This identification strategy has been used by Anelli and Peri (2019), and similar approaches have been previously adopted by Anelli et al. (2023); Hoxby (2000); Schneeweis and Zweimüller (2012); Balestra et al. (2022).

Table 2: $\text{Exp\_Dis}_{l,s,t}$ as a function of covariates and fixed effects

	(1) Exposed	(2) Exposed	(3) Exposed	(4) Exposed
School size	-0.148***			
	(0.008)			
Public school	0.651***			
	(0.007)			
Voucher school	0.367***			
	(0.007)			
Class size	0.005***	0.007***		
	(0.000)	(0.000)		
Mother vocational education	-0.031***	-0.006***	-0.003***	-0.000
	(0.002)	(0.001)	(0.001)	(0.001)
Mother college education	-0.030***	-0.003	-0.002	0.000
	(0.003)	(0.002)	(0.002)	(0.001)
Father vocational education	-0.024***	-0.005***	-0.001	-0.001
	(0.002)	(0.001)	(0.001)	(0.001)
Father college education	-0.042***	-0.008***	-0.003**	-0.002**
	(0.002)	(0.002)	(0.002)	(0.001)
Income quintile: 2	-0.031***	-0.021***	-0.002*	-0.001*
	(0.002)	(0.002)	(0.001)	(0.001)
Income quintile: 3	-0.073***	-0.024***	-0.005***	-0.000
	(0.003)	(0.002)	(0.002)	(0.001)
Income quintile: 4	-0.106***	-0.018***	-0.006***	-0.003**
	(0.004)	(0.003)	(0.002)	(0.001)
Income quintile: 5	-0.105***	-0.018***	0.002	-0.001
	(0.005)	(0.003)	(0.003)	(0.001)
Constant	0.147***	0.390***	0.608***	0.606***
	(0.011)	(0.016)	(0.002)	(0.001)
Observations	1107352	1107352	1107352	1107352
R^2	0.229	0.420	0.466	0.764
School fixed effects	No	Yes	No	No
School-letter fixed effects	No	No	Yes	No
School-cohort fixed effects	No	No	No	Yes

Notes: Standard errors clustered at the classroom level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. The table shows the association between exposure to disability and observable characteristics of schools, classes, and students (peers). Each column adds a different category of covariates and fixed effects.

6 Results

Results for Math and Spanish standardised scores are reported in Table 3 and Table 4. Having a classmate with a disability in the first grade decreases the standardized Math score by 2% to 2.5% of a standard deviation, and the effects remain quantitatively persistent over time. For Spanish, the poooled effect is of similar size, but it fades over time, from 2.6% in second grade to 1.6% by 8th grade. These results are robust to the inclusion of cohorts 2005-2009 to the analysis.[4]

The decreasing pattern could be due to distinct phenomena. Disability might have a negative effect that disappear once exposure stops, and so the pattern could be due to the higher repetition rates of students with disabilities (see Figure 1). Another possibility is the presence of learning effects, namely the gradual introduction of compensatory measures that mitigate the negative impact of exposure, whether support or better-tailored teaching methods. At present, we are not able to distinguish these two effects. However, we will be able to identify the moderating role of support measures and teacher quality with new data sources in the coming months. We interpret the vanishing effect in Spanish as evidence that the importance of cumulative learning matters: learning Math requires mastering concepts from previous grades. For Spanish, this is less important, meaning that it is easier to catch up and overcome the negative effect of exposure in initial grades.

[4]These results are not reported but available upon request.

Table 3: Effect of exposure on standardised tests scores
Cross-cohort variation

	(1) Pooled	(2) 2nd grade	(3) 4th grade	(4) 6th grade	(5) 8th grade
Panel A. Math					
Treated 1st grade	-0.013***		-0.022***	-0.020***	-0.012***
	(0.002)		(0.003)	(0.003)	(0.003)
Observations	1817441		753938	496307	567186
Panel B. Spanish					
Treated 1st grade	-0.011***	-0.025***	-0.020***	-0.009**	-0.005*
	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)
Observations	2268749	461939	750261	493562	562981

Notes: The table presents the effects of exposure on standardized math (Panel A) and Spanish (Panel B) test scores, estimated as described in Section 5. Coefficients should be interpretend as percentage changes of a standard deviation. Each coefficient comes from a different regression. *Treated 1st year* is an indicator variable that takes value one if a child had a non-repeater classmate with a disability on their first time entering primary school. All specifications control for school-letter and cohort fixed effects, as well as school size, the school share of parents with vocational and college education, class size, the classroom share of children belonging to each income quintile, and indicators for being a repeater, and own parental income and education. Additionally, specifications in column (1) include grade fixed-effects. Standard errors clustered at the classroom level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

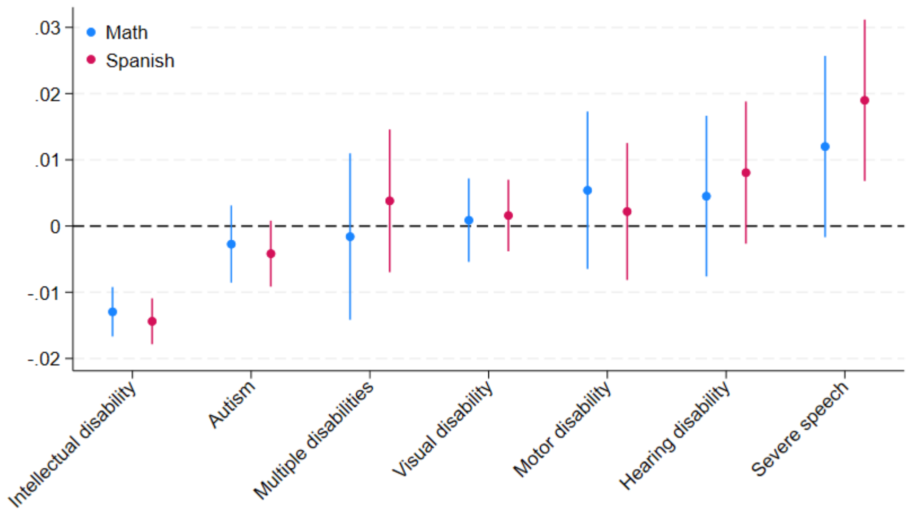
Table 4: Effect of exposure on standardised tests scores
Cross-classroom variation

	(1)	(2)	(3)	(4)	(5)
	Pooled	2nd grade	4th grade	6th grade	8th grade
Panel A. Math					
Treated 1st grade	-0.021***		-0.028***	-0.022***	-0.016***
	(0.002)		(0.003)	(0.003)	(0.003)
Observations	1817009		752095	492677	565607
Panel B. Spanish					
Treated 1st grade	-0.017***	-0.030***	-0.025***	-0.010***	-0.010***
	(0.002)	(0.005)	(0.003)	(0.004)	(0.003)
Observations	2268311	460402	748441	489930	561371

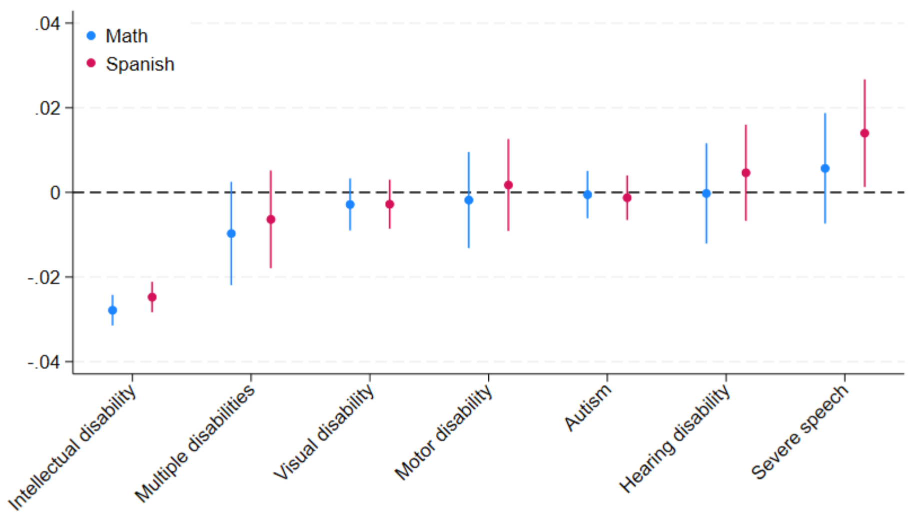
Notes: The table presents the effects of exposure on standardized math (Panel A) and Spanish (Panel B) test scores, estimated as described in Section 5. Coefficients should be interpretend as percentage changes of a standard deviation. Each coefficient comes from a different regression. *Treated 1st year* is an indicator variable that takes value one if a child had a non-repeater classmate with a disability on their first time entering primary school. All specifications control for school-cohort fixed effects, as well as school size, the school share of parents with vocational and college education, class size, the classroom share of children belonging to each income quintile, and indicators for being a repeater, and own parental income and education. Additionally, specifications in column (1) include grade fixed-effects. Standard errors clustered at the classroom level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

In Figure 6, we explore the heterogeneity of effects by type of disability. The presence of children with intellectual disabilities almost exclusively drives negative effects. Notably, the figure also shows suggestive evidence of positive spillovers of the presence of children with severe speech disorders, especially on Spanish test scores. These positive effects could be driven by increased teacher efforts, or speech-focused support measures that could have positive spillovers on peers, given the tight link with Spanish curricula in primary school. Future analysis will test this hypothesis empirically.

Figure 6: Effects by type of disability



(A) Cross-cohort variation



(B) Cross-classroom variation

Notes: The figure shows the spillover effects of disability on test scores, by type of disability. The reported coefficients were estimated as described in Section 5, but disagregating the *Expusure* dummy into several disability-type indicators. Coefficients should be interpretend as percentage changes of a standard deviation. In Panel A, we control for school-letter and cohort fixed effects, while in Panel B, we control for school-cohort fixed effects. All specifications pool all grades together and control for grade fixed effects, as well as school size, the school share of parents with vocational and college education, class size, the classroom share of children belonging to each income quintile, and indicators for being a repeater, and own parental income and education. 95% confidence bands are presented in the plot, using Standard errors clustered at the classroom level.

Measuring exposure only at the first year of each school cycle means that our causal estimand is an Intention-To-Treat (ITT) effect, namely the effect of having been exposed in the first year regardless of exposure in subsequent years. Relative to the impact of actual exposure in the year when the outcome is measured (grades 2, 4, 6 and 8 in primary school), which would correspond to a Local Average Treatment Effect estimand (LATE), the ITT could return estimates of both lower or higher magnitude.

In particular, the ITT might be lower if the negative effects of exposure disappear once exposure is lifted due to higher repetition rates for students with disabilities. In addition, there could be classes that were not exposed in the first year (classified as "untreated") that are exposed in later years due to a student with a disability from the previous cohort repeating. In future analyses, we will run robustness checks to exclude these cases and thus assess how much they impact the overall results. The ITT might also be overestimated for a selected group of classes that receive more support measures if those measures also decrease the repetition rates of students with disabilities. We are currently assembling a new dataset that includes support measures to assess their moderating effect.

Finally, note that estimating the LATE is not feasible within our context. Doing so would require estimating our regressions with an instrumental variable approach, using contemporary exposure as a treatment variable and first-grade exposure as an instrument. Nevertheless, doing so would violate two identifying assumptions. The presence of classes unexposed in the first year (not assigned to treatment) but later exposed to repeaters with disabilities (treated) would violate both monotonicity and the exclusion restriction. Therefore, any IV estimates would have no clear causal interpretation.

6.1 Moderating effects of teacher experience

In this section, we examine the moderating role of teacher experience, considering both overall teaching experience and specific experience working with students with disabilities. We then turn to the concept of teaching continuity, or "teacher looping"—the practice of having the same teacher across multiple school years, which has been shown to enhance cognitive outcomes and generate positive spillovers for classmates (e.g., Hill and Jones, 2018). We use two measures of teacher continuity. First, we analyze whether having the same teacher as in the previous year moderates the observed negative effects. While this provides some descriptive insights, the results should be interpreted with caution, as treated children are systematically more likely to be assigned the same teacher across years. To address this potential endogeneity, we use our second measure of teacher continuity: the school's historical practice of teacher looping.

Table 5 presents results from a specification that includes an interaction term with overall *Years of experience* in Equation (1). The findings indicate that each additional year of teaching experience reduces the negative effects by approximately 0.05% of a standard deviation. This implies that, to fully offset the impact of classmates with disabilites, a teacher would need around 60 years of experience. Interestingly, prior experience with students with disabilities does not mitigate the negative spillovers; in fact, it appears to exacerbate them. One possible explanation is that experienced teachers may devote a disproportionate amount of attention and resources (e.g., time, effort) to the student with a disability, potentially at the expense of the rest of the class. We will explore this hypothesis in more depth in future work.

Given the limited moderating role of general and specialized experience, we next investigate whether experience with the same group of students makes a difference. In Table 6, we interact the treatment with an indicator for having the *same teacher* as the previous year. In contrast to the results in Table 5, we find that having the same teacher across years can fully offset the negative spillovers associated with having a classmate with a disability. However, as shown in the last two rows of Table 6, treated students are about 5 percentage points (or 10% relative to untreated students) more likely to have the same teacher. This selection may upwardly bias the estimated moderating effect if teacher continuity is more common in classrooms where negative spillovers are less pronounced.

To address these endogeneity concerns, Table 7 presents a heterogeneity analysis based on the school's history of teacher looping. We compare schools without (odd-numbered columns) and with (even-numbered columns) a tradition of teacher continuity. Although the moderating effects are somewhat smaller, particularly for reading outcomes, the results still indicate that, especially in mathematics, continued teacher-student relationships can help mitigate negative spillovers.

To overcome these endogeneity concerns, in Table 7 we perform a heterogeneity analysis by separately looking at the spillover effects for schools without (odd-numbered columns) and with (even-numbered columns) a history of teacher looping. While the moderating effects are smaller than the ones before, they do show that specially for math, keeping the same teacher for more than a year can mitigate the negative spillovers. Taken together, this suggests that pursuing the policy objective of teaching continuity might not only improve overall student performance but also help in the transition from segregated to inclusive education for students with disabilities.

Table 5: Years of teaching experience effect - within cohort variation

	Math			Spanish		
	(1)	(2)	(3)	(4)	(5)	(6)
Treated 1st grade	-0.0337***	-0.0199***	-0.0286***	-0.0321***	-0.0184***	-0.0293***
	(0.0033)	(0.0026)	(0.0034)	(0.0032)	(0.0026)	(0.0034)
General experience	-0.0000		-0.0001	0.0000		0.0000
	(0.0001)		(0.0001)	(0.0001)		(0.0001)
CWD experience		0.0050***	0.0049***		0.0024*	0.0019
		(0.0013)	(0.0014)		(0.0012)	(0.0013)
Treated 1st grade × General experience	0.0005***		0.0006***	0.0007***		0.0008***
	(0.0002)		(0.0002)	(0.0002)		(0.0002)
Treated 1st grade × CWD experience		-0.0065***	-0.0075***		-0.0026*	-0.0035**
		(0.0015)	(0.0015)		(0.0013)	(0.0014)
Observations	1503740	1509325	1503740	1919901	1929102	1919901
Mean Experience	14.7508	1.2219		14.7707	1.4224	

Notes: The table presents the effects of exposure on standardized math (columns 1 to 3) and Spanish (columns 4 to 6) test scores, estimated as described in Section 5. Coefficients should be interpretend as percentage changes of a standard deviation. *Treated 1st year* is an indicator variable that takes value one if a child had a non-repeater classmate with a disability on their first time entering primary school. *General experience* corresponds to the total number of years that the students' teacher has been teaching. *CWD experience* corresponds to the years of previous experience that the teacher has had with students with disabilities. Each column corresponds to a different specification were all grades are polled together. All specifications control for school-cohort fixed effects and grade fixed effects, as well as school size, the school share of parents with vocational and college education, class size, the classroom share of children belonging to each income quintile, and indicators for being a repeater, and own parental income and education. Standard errors clustered at the classroom level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 6: Same teacher as the year before - Across cohort variation

	(1) Pooled	(2) 2nd grade	(3) 4th grade	(4) 6th grade	(5) 8th grade
Panel A. Math					
Treated 1st grade	-0.031***		-0.041***	-0.026***	-0.025***
	(0.003)		(0.004)	(0.004)	(0.004)
Same teacher as t-1	0.049***		0.058***	0.053***	0.035***
	(0.003)		(0.005)	(0.005)	(0.004)
Treated × Same teacher	0.035***		0.039***	0.014**	0.025***
	(0.004)		(0.006)	(0.006)	(0.005)
Observations	1671273		696527	454540	520192
Panel B. Spanish					
Treated 1st grade	-0.021***	-0.041***	-0.032***	-0.011**	-0.009**
	(0.002)	(0.006)	(0.004)	(0.005)	(0.004)
Same teacher as t-1	0.035***	0.030***	0.036***	0.034***	0.035***
	(0.003)	(0.006)	(0.005)	(0.005)	(0.005)
Treated × Same teacher	0.019***	0.024***	0.025***	0.006	0.007
	(0.003)	(0.007)	(0.006)	(0.006)	(0.006)
Observations	2072815	417153	692440	452161	511047
% same teacher (Not treated)	51	53	52	48	51
% same teacher (Treated)	56	60	57	49	56

Notes: The table presents the effects of exposure on standardized math (Panel A) and Spanish (Panel B) test scores, estimated as described in Section 5. Coefficients should be interpretend as percentage changes of a standard deviation. *Treated 1st year* is an indicator variable that takes value one if a child had a non-repeater classmate with a disability on their first time entering primary school. *Same teacher as t-1* corresponds to an indicator for students who have the same teacher as in the year prior to the exam. Each column corresponds to a different specification. All specifications control for school-leter and cohort fixed effects, as well as school size, the school share of parents with vocational and college education, class size, the classroom share of children belonging to each income quintile, and indicators for being a repeater, and own parental income and education. Additionally, specifications in column (1) include grade fixed-effects. Standard errors clustered at the classroom level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 7: Heterogeneity by teacher looping - Across cohort variation

	Pooled		2nd grade		4th grade		6th grade		8th grade	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A. Math										
Treated 1st grade	-0.018***	-0.010***			-0.028***	-0.007	-0.023***	-0.013**	-0.015***	-0.005
	(0.003)	(0.002)			(0.004)	(0.006)	(0.004)	(0.007)	(0.003)	(0.005)
Observations	535958	1281483			511060	242878	380654	115653	391630	175556
Panel B. Spanish										
Treated 1st grade	-0.013***	-0.011***	-0.023***	-0.028***	-0.023***	-0.015***	-0.012***	0.004	-0.004	-0.007
	(0.003)	(0.002)	(0.006)	(0.006)	(0.004)	(0.005)	(0.004)	(0.007)	(0.004)	(0.006)
Observations	641593	1627156	241199	220740	479942	270319	386712	106850	408682	154299
Looping	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: The table presents the effects of exposure on standardized math (Panel A) and Spanish (Panel B) test scores, estimated as described in Section 5. Odd-numbered columns are run on the subsample of schools that do not have a history of teacher looping. Even-numbered columns are run on the subsample of schools that do have a history of teacher looping. Coefficients should be interpretend as percentage changes of a standard deviation. *Treated 1st year* is an indicator variable that takes value one if a child had a non-repeater classmate with a disability on their first time entering primary school. Each coefficient comes from a different regression. All specifications control for school-leter and cohort fixed effects, as well as school size, the school share of parents with vocational and college education, class size, the classroom share of children belonging to each income quintile, and indicators for being a repeater, and own parental income and education. Additionally, specifications in columns 1 and 2 include grade fixed-effects. Standard errors clustered at the classroom level, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

7 Future extensions and conclusions

As an increasing number of countries move towards the inclusion of students with disabilities into mainstream education, understanding the consequences for peers and the best implementation strategies becomes a priority to policymakers.

In this paper, we estimate the causal effect of exposure to classmates with disability on peers' grades in primary school We rely on rich administrative data covering the universe of students from Chile, a country that implemented a series of inclusion reforms starting in 2009. The large sample size allows us to overcome the issue of small and selected samples that are often found in studies on severe disabilities, making it possible to estimate causal effects that are so far unknown.

Results indicate the presence of negative spillovers on peers' grades that are persistent for math but vanish over time for Spanish. We attribute this finding to the importance of cumulative learning in Math. Having the same teacher across different grades, especially in the first years of primary education, appears to be a highly effective countermeasure against negative spillovers, resulting even in the reversal to positive spillover effects in the case of Math. Future analyses from ongoing work will focus on additional survey data to disentangle the specific role of classroom environment.

These results underline both the challenges related with inclusive education and the policies that can be enacted to face them. Our results indicate that, while exposure to disability has an adverse effect on peers' schooling outcomes, a relatively easy and unfocused policy intervention aimed at ensuring teacher continuity is able to fully offset these effects. Teacher continuity has been shown to benefit students' cognitive and non-cognitive outcomes regardless of exposure to disability: it is already recommended, and our results indicate large positive spillovers in the context of inclusive education. We also provide suggestive evidence that targeted support measures can also have positive spillovers on the test scores of classmates without disabilities. Our ongoing work will ascertain whether these measures translate in an improved class environment, which could act as a mediator. Already at the current stage, however, our results indicate that a large-scope and relatively inexpensive policy aimed at teacher continuity, that does not require specific skills or experiences, could represent a viable strategy to prevent potential negative spillovers.

References

M. Anelli and G. Peri. The effects of high school peers' gender on college major, college performance and income. *The Economic Journal*, 129(618):553–602, 2019.

M. Anelli, K. Shih, and K. Williams. Foreign students in college and the supply of STEM graduates. *Journal of Labor Economics*, 41(2):511–563, 2023.

S. Balestra, B. Eugster, and H. Liebert. Peers with special needs: Effects and policies. *Review of Economics and Statistics*, 104(3):602–618, 2022.

A. Carrasco, G. Gutiérrez, and C. Flores. Failed regulations and school composition: selective admission practices in chilean primary schools. *Journal of Education Policy*, 32(5):642–672, 2017.

S. E. Carrell and M. Hoekstra. Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics*, 2(1):211–228, 2010.

S. E. Carrell, M. Hoekstra, and E. Kuka. The long-run effects of disruptive peers. *American Economic Review*, 108(11):3377–3415, 2018.

M. P. Godoy, M. L. Meza, A. Salazar, et al. Antecedentes históricos, presente y futuro de la educación especial en chile. 2004.

A. J. Hill and D. B. Jones. A teacher who knows me: The academic benefits of repeat student-teacher matches. *Economics of Education Review*, 64:1–12, 2018.

C. Hoxby. The effects of class size and composition on student achievement: new evidence from natural population variation. *Quarterly Journal of Economics*, 115(4):1239–1285, 2000.

A. e. a. Kalambouka. The impact of placing pupils with special educational needs in mainstream schools on the achievement of their peers. *Educational Research*, 49(4):365–382, 2007.

E. A. Nowicki and R. Sandieson. A meta-analysis of school-age children's attitudes towards persons with physical or intellectual disabilities. *International Journal of Disability, Development and Education*, 49(3):243–265, 2002.

N. M. Ruijs, I. Van der Veen, and T. T. Peetsma. Inclusive education and students without special educational needs. *Educational Research*, 52(4):351–390, 2010.

N. Schneeweis and M. Zweimüller. Girls, girls, girls: gender composition and female school choice. *Economics of Education Review*, 31(4):482–500, 2012.

J. P. Valenzuela Barros and G. Toledo Román. Ordenamiento de profesores y estudiantes entre y dentro de los establecimientos escolares: el caso de chile. 2012.

Acknowledgements

Nicoletta Balbo and Alice Dominici acknowledge the financial support received by the European Union (ERC FRAILIFE—Child Disability and Family Life, G.A. 101077533) under the European Union's Horizon 2021-2027 research and innovation program. Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.