The effect of the Integration of Students with Special Educational Needs:

Evidence from Chile

Abstract

This paper studies the effect of the integration of students with Special Educational Needs (SEN) on the academic achievement of their peers without SEN. This achievement is measured using performance in standardized reading and mathematics tests. The study also evaluates the effect of a policy that recognizes and increases resources for special educational needs that had not yet been considered and improves education provision protocols for students with SEN. Using administrative data and standardized test scores, we constructed a panel that follows a cohort of students before and after the reform, determining for each individual and year whether the class to which he/she belongs has any students with SEN. Our identification strategy employs panel data with fixed effects at the school, individual, and time levels. Estimates show that, on average, having a peer with SEN in the classroom has a negative effect on the academic performance of students without SEN. However, these effects are small and decrease, or even vanish, once better inclusion policies are in place. These results suggest that the effect on peers is almost totally canceled when more resources are provided and when adequate treatment and support protocols are implemented.

Keywords: special educational needs, inclusive education, public policy

JEL Classification: I21, I28, I18, J14

1. Introduction

Over the last years, inclusive education and the effects of diverse classrooms on students have attracted growing interest in the economic literature (for instance Angrist & Lang, 2004; Gould et al., 2009; Raitano et al., 2011; Sacerdote, 2011). However, quantitative studies analyzing the integration of students with special educational needs (SEN) in regular classrooms[1] are scarce and mostly incipient (for instance, Jordan et al., 2009; Keslair et al., 2012; Ruijs, 2017; Stiefel et al., 2019; Gottfried & Kirksey, 2019). The lack of available evidence is probably due to the limited amount of suitable information and the natural difficulties associated with identifying comparable groups.

It is open to debate whether the inclusion of students with SEN in the regular education system is desirable (see for instance, Ruijs & Peetsma., 2009; Gottfried et al., 2016; Stiefel et al., 2019). The arguments in favor of their inclusion are that the SEN students would develop to their full potential in an integrated classroom, and the non-SEN student would develop skills such as tolerance, and patience. The argument against is that inclusion may have a negative impact on non-SEN students because they would be distracted or would receive less attention from their teachers due to the presence of SEN students.

Studies on this topic have mainly been conducted in developed countries and emphasize the effect of this approach on students with SEN. Three elements are key for this situation. First, the availability of abundant information, captured in a systematic manner. Second, the increasing interest attracted by inclusive policies over the last years. Third, a larger amount of resources enabling developed countries to implement these policies. The literature reports mixed findings with respect to the academic outcomes of integrated students (For instance, Salend et. al., 1999; Stiefel et al., 2019; Gottfried et al., 2019). Regarding non-integrated students, research is even scarcer and mixed. The findings reveals a null (Aizer, 2008; Friesen et al., 2010; Ruijs, 2017), small negative (Fletcher, 2009; Kristoffersen et al., 2015), or positive impact in academic performance (Hanushek, 2002).

In this paper, we analyze the effect of the inclusion of students with SEN on the academic achievement of their peers without SEN and how it changes after a policy change that occurs in 2010. Should be considered that the people of interest in this paper are the non-SEN students that share classroom with SEN students, and no directly the SEN students[2]. The policy was implemented in Chile and consists in a change of the treatment of students with SEN.

This article contributes to the existing literature in at least three areas. First, it provides evidence for a developing country (Chile). There have been few studies examining this issue outside of the United States with the quality of data employed in the current study. Differences in institutional frameworks, more limited resources, and poorer academic performance, among other elements, suggest that results may not be comparable to developed countries. Second, the article provides new

[1] The term "regular classrooms" will be used to refer to those not only devoted to educating students with SEN, in contrast with those specifically focused on them.

[2] There is no feasible and comparable information about the performance of students with SEN; therefore, we are unable to address the effects of the policy on them in this paper.

evidence of the effects of the inclusion of students with SEN on their peers without SEN in terms of academic performance differentiating results for students with temporary SEN (henceforth TSEN) and permanent SEN (henceforth PSEN). Third, it examines the impact of a specific policy and provides insight into improving inclusion such that all children benefit. Indeed, we find evidence showing that the apparent negative effects associated with the integration of students in the classroom dissipate when inclusive policies are implemented. These include additional resources, more availability of specialized professionals, and support measures for integrated students.

In order to do this, we took advantage of the policy change implemented in Chile in 2010, through which TSEN were recognized and the protocol for treating students integrated into regular classrooms was formalized. The recognition of TSEN led to the creation of a treatment and education protocol and an increase in the resources allocated to schools in connection with these students. Prior to the policy, students with TSEN were not recognized as such and therefore received exactly the same resources as students without SEN. On the other hand, before the policy, students with PSEN did receive additional resources, but there was no clear integration and treatment protocol in place. For those with PSEN, the policy led to improved treatment and education protocols.

We use unique administrative databases that make it possible to identify students with SEN who are placed in regular classrooms. Indeed, we are able to determine whether a student has SEN during a given period. In addition, databases published in connection with the SIMCE standardized tests enable us to identify the scores of students without SEN during the same periods. Then, using SIMCE test data, we follow a student cohort for several years. That data enables us to construct a panel that contains a student's scores, specifies whether he/she has classmates with SEN, and includes his/her socioeconomic information. We selected the cohort of students who took the SIMCE in 2007 and 2011 (fourth and eighth grade respectively). Thus, using a fixed effects model at the school, student, and time levels, and controlling for observable characteristics, we estimated the impact of the policy change and the inclusion of students with SEN (integrated students) in regular classrooms on academic achievement of students without SEN (non-integrated students). In general we argue that with the fixed effect model it is possible to reduce different bias such as individual specific characteristics, school policies and time trends. The Empirical Methodology section provides an in-depth description of this approach, the problems addressed with it, the main assumptions employed to interpret data, and the robustness tests conducted.

Results show that, during the first period, having a classmate with SEN is associated with a small negative effect on standardized test scores. During the second period, when the policy change (treatment protocol and additional resources) began to operate, negative results in standardized test decreased or became slightly positive.

The rest of the paper is organized as follows. Section 2 summarizes the main results reported in the literature. Section 3 describes the Chilean educational context, the treatment of students with SEN, and the policy change in 2010. Sections 4 and 5 describe the data and the identification strategy employed in this paper. We present our main results in section 6. Finally, section 7 presents our conclusions and future challenges that we think should be explored.

2. Literature review

The inclusion of SEN students have been generally analyzed in the literature from one of three main perspectives. First, from the SEN students perspective (For instance Stiefel et al., 2019; Gottfried et al., 2019). Second, from the non-SEN students that share classroom with SEN students perspective (For instance, Hanushek et al., 2002 ; Azier, 2008; Fletcher, 2009; Gotfried, 2014). And third and less common, from the teacher or school provision perspective (For instance, Jordan et al., 2009). In this paper we are mainly focus in the second perspective, though we use a variation in a policy that defines how schools provide this education. Should be noticed that due to data restrictions, we are unable to see academic outputs of SEN students. Therefore, we are not able to study the first perspective.

We argue that our main contributions are that (i) there have been few studies examining this issue outside of the United States with the quality of data employed in the current study; (ii) it presents detailed results considering the type of SEN (Permanent or Transitory); (iii) and it examines the impact of a specific policy and provides insight into improving inclusion such that all children benefit.

The studies that examine how the inclusion of SEN students in regular classrooms impacts the academic performance of students without SEN have reported mixed results and are mainly located in developed countries.

Hanushek et al. (2002) is one of the first systematic studies on this topic. They studied the integration of students with SEN in Texas. The authors estimates using panel data controlling for time, student, and school-grade fixed effects. Their estimates show that the classmates of integrated students improve their academic performance, with the impact positively correlating with the percentage of students with a disability in the classroom. The authors warn that there is a positive association between special needs education and academic achievement, but that it is not possible to discern if it is caused by the larger amount of resources allocated due to the higher rate of integrated students or by changes in the classroom that they failed to consider.

Fletcher (2009) analyzes how the inclusion of students with emotional and behavioral disorders (EBD) in regular classrooms impacts academic achievement. He analyzes the US case employing OLS estimation with repeated cross-sectional data to control for school fixed effects. Results show that the impact on classmates without SEN is negative but small in reading and null in mathematics.

Aizer (2008), also for the U.S., examines the effect of an attention deficit hyperactivity disorder (ADHD). Before diagnosis, this author's estimates reveal a negative impact on academic achievement for students who share a classroom with students with ADHD. However, the negative effect disappears once the disorder is identified and treated. Many prior studies focus on the academic benefits of separating disruptive students. This article shows that it is not necessary to wholly rearrange the classroom to mitigate potential negative externalities associated with inclusion if policy frameworks provide suitable tools to assist the children who require support.

Gottfried (2014), using US data, conducted a study examining the effect of inclusion on non-cognitive skills. This author used longitudinal data to implement fixed effects by school and by

year. The article makes main contributions are that it demonstrates a negative impact on non-cognitive skills; it identifies effects associated with TSEN and PSEN separately; and it identifies contextual factors (specifically, teacher experience) that mitigate the results observed, which should be taken into account when designing public policies[3]. In the same line and using the same data, Gottfried et al. (2016) analyze if the ED students integration have an effect over absences. They found that the students in classrooms with ED students are more likely to be absent than they peers. Should be noticed that, a difference from those studies, in our paper we are studying academic achievement (and no non-cognitive skills).

Friesen et al. (2010) study the effects of inclusion in Canada. Their methodology is mainly based on longitudinal data for multiple student cohorts in British Columbia. They use this data to estimate fixed effects at the school and grade level. The authors' estimates suggest that sharing a classroom with same-grade peers who have learning and behavior problems have no impacts on academic outcomes.

Recently, and more similar to our paper, Ruijs (2017) study the effect of SEN students over the classmates in Netherlands. His empirical strategy uses three independent identification approaches: fixed effect at individual level, fixed effects at school level, and a neighborhood variation. He finds that the special needs students do not have a statistically significant effect on the academic achievement of their classmates. However the empirical strategy could be similar[4], our paper differentiate from Rujis (2017) because of three main reasons. First, we are able to analyze a developing country as Chile that have a more limited budget, different institutional frameworks, and poorer academic performance, among other elements. This different setting may produce different results from those found for developed countries. But also, Chilean context allows us to have more detailed data of SEN students. Second, we are able to analyze a policy change that show us that the non-effect result could be driven by extra resources, suggested by his work, and the accurate attention. In that line, our work is able to reconcile the non-effect and negative effects literature. And third, our data allow us to differentiate between permanent and transitory needs, showing heterogeneous effects between both types of SEN.

3. The Chilean context

3.1. Institutional framework

In the period analyzed, the Chilean system was composed of preschool education; primary education (first 8 grades); secondary education (next 4 grades, from ninth to twelfth grade); and higher education[5]. Students can choose to attend one of three types of centers: private schools,

[3] Considering the mixed evidence available, Gottfried (2014) opines that the integration of students with SEN in regular classrooms affects performance through direct channels, linked to student interaction in the classroom, and through indirect channels, related to the inclusive education provided by the system. The final result will depend on the magnitude of the effects associated with each channel. Regarding direct channels, interaction in a diverse classroom can generate (i) positive effects, resulting from the potential development of interpersonal skills, and (ii) negative effects, linked to a potential increase in disruptive behavior in the classroom. With respect to indirect channels, the inclusive approach can generate (i) positive effects, resulting from resource reallocations or increases benefiting inclusive education for students with and without SEN, and (ii) negative effects, connected to distortions within the classroom or reduced teacher attention toward students without SEN, or both.

[4] We use an individual and school fixed effect opposite to Rujis (2017) that use three different approaches that includes both kinds of fixed effects.

[5] Primary and secondary education (12 years) are mandatory in Chile.

State-subsidized private schools (henceforth voucher schools) and State schools. These institutions served, respectively, 7%, 48%, and 45% of the total student population in 2011. Private schools receive no State funding and depend on fees paid by students' families. State schools receive public funding, do not charge families, and are administered by the local municipality. Voucher schools are funded partly by the State and partly by families and are administered by private parties. These State-funded schools receive equal subsidies for each student enrolled (a demand subsidy).[6] Schools commonly have more than one class per level in order to cover demand and benefit from economies of scale given the number of students that they serve[7].

Apart from the classification by type of funding, the system has schools that provide regular education and others that offer special education. Special education schools are specifically aimed toward students with SEN, whereas regular schools serve students with and without SEN. This means that students with SEN can attend both special and regular schools. According to the definition of SEN published by the Ministry of Education after the policy change, PSEN include hearing, visual, intellectual, and multiple disabilities (more than a single type), autism, dysphasia, and deafness-blindness, whereas TSEN include attention deficit hyperactivity disorder, specific language disorders, and borderline performance in IQ tests with significant limitations in adaptive behavior or borderline intellectual functioning. Before the change, SEN were only those later defined as PSEN.

In general, students with SEN who attend regular schools do so via School Integration Projects (SIP). These projects were developed by the State in order to integrate students with SEN in regular schools. In 2013, students with SEN represented 6.4% of the population of regular schools[8]. That year, according to Ministry of Education data, 45% of students in regular schools had at least one classmate with SEN. Schools that receive State funding can request to implement a SIP voluntarily. To apply, the school must design a program containing: (i) support staff; (ii) professionals who will diagnose SEN, who must be authorized by the Ministry and be on its listing of special education professionals; (iii) students who will take part in the SIP; and (iv) a program in which competent professionals define support measures for students in regular classrooms and plan interventions, evaluations, collaborative work, and coordination among teachers, as well as specific individual tasks for students with SEN. The Ministry of Education is responsible for evaluating and approving the SIP. Schools that develop a SIP receive an additional subsidy for each student with a treated SEN (the following subsection details the amounts provided per integrated student). The Ministry of Education supervises and regulates the use of additional resources[9].

[6] According to Ministry of Education data for 2002, 90% of voucher schools received copayments from families. This limits access to such schools for many families. In addition, private and voucher schools can select their students according to self-defined criteria. According to Contreras et al. (2010), 6% of public schools and 56% of voucher schools engaged in some type of student selection process in 2005. Voucher schools can select students based on academic examinations, interviews with parents, and specific educational projects (religious schools).

[7] Each school defines the number of places that it offers according to the number of students per class and the resources available to each level (e.g. a school can have three fourth grade classes).

[8] This percentage is lower than those reported in other countries. Some students with SEN attend special schools (4.73% of all students in the Chilean educational system), but others have not yet been identified, although they are served by regular schools.

[9] Schools with integration projects can use these additional SEN resources for the following purposes: hiring specialized support professionals, acquiring specific teaching materials, implementing teacher training programs, performing infrastructure adjustments, training staff, and compensating teachers and administrators for the time devoted to collaborative projects and

SEN must be diagnosed by a team of professionals certified by the Ministry of Education. These professionals have been trained to perform this task and are on the Ministry's listing of authorized professionals. Evaluation is based not only on the identification of diseases, disorders, and other problems but also on the student's interaction with his/her educational environment, particularly with teachers and family. This diagnosis is based on criteria set out by WHO in 2001, which has defined international standards.

A point of concern is that schools could have started offering incentives to encourage positive SEN diagnoses after the policy was implemented. We believe that four factors mitigate these possible incentives. For several reasons, the mere existence of the policy is unlikely to have resulted in more reported cases of SEN. First, the Ministry of Education is rather strict with the identification of students with SEN, thus hindering collusion between schools and certifiers leading to the misdiagnosis of SEN. Second, although students with TSEN were first recognized in 2011, given our methodology, they are exactly the same as those we acknowledged in 2007 (in this regard, we did not observe a substantial increase in their numbers across years). Third, schools have monetary incentives to report SEN when they do exist (in the case of PSEN, the incentive exists in both periods), which makes it unlikely for such students to be underreported. Fourth, the monetary incentives for schools to enroll students with PSEN remained stable across periods and became subjected to more stringent regulations; therefore, schools are unlikely to overreport the number of said students.

3.2. *Policy change*

In 2010, a new inclusion policy was introduced, reforming the treatment of students with SEN. In general terms, the policy involved (i) a broader definition of SEN, which led to the recognition and treatment of students with such needs, as well as the allocation of additional resources to them; and (ii) the establishment of new protocols for treating and educating students with SEN, stricter and more specific than existing ones in the case of PSEN and novel in the case of TSEN. The following subsection details the policy change in terms of the resources allocated and the new student treatment protocols.

3.2.1. *Additional resources for integrated students*

Students with PSEN received US$286 per month before and after the reform. This total comprises the regular subsidy (US$94, the baseline amount received by all students) and a special subsidy (US$192). Students with TSEN, who were not recognized as such, received the regular monthly subsidy (US$94) before the reform. After the implementation of the new policy, TSEN students were granted an additional subsidy of US$155, which increased their monthly subsidy to US$249. It should be noted that the increase in funding allocated to schools and the investments associated with integration projects could also benefit students without SEN (for instance, thanks to the services provided by new professionals in schools or through the development of techniques aimed at improving concentration in the classroom).

program evaluation duties. These schools can integrate a maximum of two students with PSEN and five with TSEN in a single classroom.

3.2.2. Treatment of students with SEN

Before 2010, the law defined no specific requirements for the treatment of students with specific SEN in the classroom. The team of independent professionals in charge of diagnosing SEN was also tasked with designing support measures for each student. The main characteristic of the integration program used before the policy change was the "resource room" where students with SEN were taught separately from the rest of the school, receiving support and care. Depending on their diagnosis, students could attend a regular classroom only, both a regular classroom and the resource room, or the resource room only. In the latter case, they were able to join their peers in the regular classroom for special events or recreational activities.

After 2010, the law provided specific instructions for addressing each of the recognized special educational needs. A student with SEN who attends a school where a SIP has been implemented can only be taught in a regular classroom, due to the elimination of the resource room[10]. The law also provided more specific information about the minimum components of the assistance for each identified SEN.

4. Data

Two data sources were used in this study. First, administrative databases, which enable us to identify students with SEN; second, SIMCE databases, which allow us to observe a student's individual scores in standardized tests and some of his/her observable characteristics.

Private schools were excluded from our analysis because they (i) represent a very small part of the total number of students in Chile and (ii) serve only a minor number of students with SEN[11].

4.1. Identification of SEN

The Ministry of Education's administrative databases have included information about students diagnosed with SEN since 2011. In 2013, they start including detailed information whether students display PSEN or TSEN. Thus, it was possible to identify the students who had a recognized SEN diagnosis in 2011, but not if their specific needs were permanent or temporary. Therefore, we identified SEN in each period through inference based on 2013 data[12].

For the inference, we adopt two different approaches for each type of SEN. In the case of the permanent needs, we assumed that a student with PSEN in 2013 would also have PSEN in 2011 and 2007 -considering the stable characteristics of permanent special educational needs-.[13] In the

[10] This means that there is not "resource room" after the policy change.
[11] In 2011, 7% of students were enrolled in private schools in Chile. 0.31% of them were integrated students.
[12] The identification of special educational needs for 2007 is based on a database of school self-reports (less reliable). For this reason, 2007 data are less valid than 2011 and 2013 data. The latter are derived from Ministry of Education administrative records, collected through evaluations conducted by a multidisciplinary team of professionals who are registered and approved by the authorities. Therefore, the 2007 SEN database was neither used nor alluded to in this version of the article. The 2007 SIMCE databases are used instead, as explained in the following paragraphs.
[13] Alternatively, we could have established that students with TSEN are those defined as such in the 2013 administrative database. This approach would have involved assuming that TSEN affects students permanently, without considering the temporary component of such needs. To differentiate both approaches, we compared potential discrepancies between them. Table A.1 (in the appendixes) details our comparison between these identification approaches. Approximately 27% of the students who we identified as having TSEN are also recognized as such in 2013. This group was labeled (c) and is validated in both databases. Group (b) (35%)

case of the transitory needs, we recognized as students with TSEN those who, having been identified as SEN in 2011, were not recognized as PSEN[14].

Approximately 3% of students had SEN (temporary or permanent) in both years. A 1% had PSEN and a 2% had TSEN. In 2011, 3% of the sample had SEN (1% PSEN and 2% TSEN). 69% of these students attend public schools, 31% attend voucher schools, and only 0.2% attend private schools. In 2007, 4% of the sample had SEN (1% PSEN and 3% TSEN). Of these students, 76% attended public schools, 23% attended voucher schools, and only 0.4% attended private schools.

4.2. *SIMCE scores and observable characteristics*

In order to construct a panel of the academic performance of students before and after the reform, we used mathematics and reading SIMCE scores for the cohort of students that we observed in 2007 (fourth grade) and 2011 (eighth grade)[15]. SIMCE data includes standardized test results and also a set of questionnaires that enable us to characterize schools, families, and students[16].

In this paper, we use SIMCE test scores as a measure of students academic performance. SIMCE is a national test and it is taken by all the students that are in the specifics grades evaluated[17]. The SEN students are allowed to skip the test, and their score is not considered in the school achievement. Therefore, we are not able to analyze the academic performance -measured as the SIMCE score- of SEN students.

4.3. *Analytic Sample*

Table A.2 presents, for the initial cross-section samples, the total number of students with SEN and the percentage that they represent, by school type. 3.3% of the total number of students had SEN in 2007, whereas 3.8% had SEN in 2011[18]. This increase can be almost fully explained by the influence of students with SEN who attend municipal schools, who rose from 4.5% to 5.6% between 2007 and 2011.

Similarly, Table A.3 analyze students without SEN[19]. Regarding students without SEN, 2007 data shows that the number is nearly identical across school types (public and voucher). Nevertheless,

---

comprises students identified as having TSEN according to our methodology, but not according to the approach adopted in the 2013 database. Presumably, these individuals, diagnosed with TSEN in 2011, are not classed as special needs students in the 2013 database because they have overcome their difficulties and are no longer recognized as such by educational authorities. Finally, group (a) (38%) is composed of students not identified as having TSEN according to our methodology, but who are identified as such in the 2013 database. These are students who received a late diagnosis or whose special needs emerged after the year 2011. In both cases, this group is rightly identified as non-TSEN.

[14] It should be noted that some students with SEN may have been neither identified nor treated through the educational policy in place. Thus, in this study, the effects found and the results discussed are only those associated with the integration of students with SEN recognized by this methodology. Studies conducted outside of Chile report that 1 out of 5 children have undiagnosed SEN (Warnock Report, 1978).

[15] In general, both grades have a similar grade structure. The students are in one class (this is, one group of students) throughout the day and they could rotate of classroom based on the subject.

[16] Thus, we obtained information about students' age and gender, household income, number of books in the home, parents' education level, the administrative region of each school, their location (rural or urban), and its type (State, voucher, or private).

[17] In each year are evaluated specific grades. In general 4th, 8th and 10th grade are included.

[18] Specifically, it is PSEN that are behind this increase, given the rising number of students with such needs. Regarding the PSEN - TSEN ratio, in both years the former account for nearly 1/3 of the total, while the latter comprise nearly 2/3.

[19] The percentages shown are relative to each category and not to the total number of students.

the number of students who attended voucher schools in 2011 rose[20]. In contrast, students without SEN present in both periods a municipal-voucher school ratio that closely resembles that of the full sample. This is true of all integration categories, which suggests that, despite the loss of observations in the panels, the samples represent the original composition of each period. Finally, the number of students without SEN observed in both periods amount to 120,664, representing 70% of the total number of students without SEN in 2011.[21]

Table 1 shows descriptive statistical values for the cross-section and panel samples. To rule out selection bias derived from the loss of these observations, we compared the means of the main variables for the cross-section and panel samples. Regarding integration, the share of students with at least one SEN classmate, the share of students with at least one TSEN classmate, and the share of students with at least one PSEN classmate are statistically equivalent both for the cross-section sample and the panel data. Similarly, the share of students who attend voucher schools is statistically equal in all samples, reaching approximately 50%. Finally, average test scores are similar for the panel sample than for the cross-section samples. In consequence, no relevant bias is observed when inspecting the data. It could be stated that the population in the panel sample has better academic performance than the population in the initial sample. This is because students who repeat a grade or who are at risk of dropping out of the system are the lowest performing.[22].

Table A.4 clarifies the variations in the share of students with integrated classmates. This table presents information for the samples that make up the 2007-2011 panel. Panel (A) shows that the number of classes in the system decreased during the period studied[23]. Also, this panel shows that 38% of classes integrate at least one student with SEN, 15% integrate at least one student with PSEN, and 29% integrate at least one with TSEN. These levels remain stable in 2011, except for the percentage of classes that integrate PSEN, which rises to 21%[24]. On the other hand, in the panel (B) could be observed that in 2007 it amounts to 15.3 students per class. While the average number

[20] This growth is not due to an actual increase in enrollment rates; instead, the database for the year 2007 lost a large number of students from these schools due to invalid observations (which lacked SIMCE scores or information about the characteristics vector). Since we have no information about the socioeconomic characteristics of these observations, it is not possible to determine whether this attrition concerns a specific group of students from voucher schools.

[21] Tables A.2 and A.3 show that students with SEN make up a small part of the total number of students, most of whom attend municipal schools. In this group, most students have TSEN. The total number of students with SEN rose during the 2007-2011 period, with the PSEN group displaying the largest increase. In addition, the percentage of students with SEN who are not observed in both periods is significantly bigger than that of students without SEN. Students with PSEN comprise the smallest group in both periods, which probably stems from the learning difficulties associated with the nature of their needs.

[22] If these individuals –who are not considered– perceive an effect due to the integration of students with SEN that differs from that perceived by the rest of the population, their exclusion will introduce bias into the estimation of the coefficients of interest. However, this bias is not easy to determine: it is not obvious that lower-performing students are more strongly affected than the rest of their classmates, nor is it clear that the effect of integration is positive or negative for them. Let us consider the following example. Imagine a student population composed of two groups –high level and low level– with the former being those observed in the period studied and the latter being those excluded due to dropout or grade retention. The real effect of integration on the academic performance of each group is Eh and El, both negative. If $|Eh|<|El|$,, the effect calculated with our methodology Ee will be an overestimation of the real effect of the whole population Er, i.e. Ee>Er. In contrast, if $|Eh|>|El|$, the effect is underestimated, i.e. Ee<Er. It is not trivial to determine which case applies. The first case is valid when low level students are more academically affected by eventual disruptive behavior in the classroom or by reduced attention from teachers. However, these students –at one end of the spectrum– may perform so poorly that they may not be reactive to a worsening of the conditions in which classes take place. In these circumstances, the second case applies. Lastly, the assumption that Eh and El are negative is not evident either.

[23] This is manifested through an increase in the average size of the classes without SEN integration –nearly two additional students per classroom– (see first row, panel B).

[24] It should be noted that the number of students identified as having PSEN and TSEN increases during this period, especially in the case of the PSEN case.

of students in classes without SEN integration increase in 2011, classes with SEN integration lose about two students on average[25]. The reduction in the average size of integrated classes suggests that schools may have structured them in a way that increases teachers' ability to respond to the special needs of integrated students. Since this would be a variation in school behavior over time, it is not captured by fixed effects. In consequence, all our estimations also control for class size.

5. Empirical strategy

5.1. Identification problems

Estimating the impact of integration on academic achievement requires solving several issues that can hinder proper identification. We detail these problems in the Appendixes section (A1 appendix), and discuss how this study aims to solve them or mitigate their effect on the estimation process.

5.2. Empirical strategy

This study estimates the effect that being educated in a classroom with at least one student with SEN has on the academic performance of classmates without SEN. This analysis differs depending on whether the classroom integrates students with TSEN or PSEN and considering the period examined (before or after the reform). We use longitudinal data to estimate fixed effects.

For the reasons previously stated (see A1 appendix), our preferred specification includes individual-level, school-level and year fixed effects[26]. The equation to be estimated[27] is the following:

$$Y_{ijt} = \alpha_i + \lambda_t + \phi_j + \beta_1 * TSEN_{it} + \beta_2 * PSEN_{it} + \beta_3 * TSEN_{it} * TIME + \beta_4 * PSEN_{it}ME + \gamma * X_{itj} + \varepsilon_{ijt} \qquad (1)$$

Where Y_{ijt} is the score of student i attending school j during period t, α_i, ϕ_j and λ_t are individual, school and time fixed effects, respectively. PSEN and TSEN are dummy variables that adopt the value 1 if the student "i" attending school "j" during the year "t" has at least one classmate with PSEN or TSEN, respectively. Time is a dummy variable with value 1 in 2011 and 0 in 2007. Lastly, X_{ijt} control for student characteristics, family, class, and the school attended by subject "i" during the year "t".

Parameter β_1 is the estimation of the effect of TSEN inclusion conditional to the first period (2007) and $\beta_1 + \beta_3$ is the effect of TSEN inclusion conditional to the second period. β_1 can be interpreted as the effect of TSEN inclusion when these students were neither recognized nor assisted by the system. $\beta_1 + \beta_3$ can be interpreted as the effect of TSEN when the new policy is implemented. β_3 is the effect of the

[25] The largest drop affects classes with TSEN integration, where the average number of students decreases by approximately two per class. This largely explains the decrease in the percentage of students with a classmate with TSEN (table 1). The size reduction of the classes with PSEN integration is smaller. In addition, the increase in the percentage of these classes (third row, panel (A)) and the larger number of PSEN students identified (table A.3) account for the larger percentage of students with PSEN classmates (table 1).

[26] Alternatively, we estimated a model with individual level and time fixed effects, but without school level fixed effects. The specification allows us to control by two sources of bias. First, parents' preferences for inclusive education, and their potential relationship with academic results. Second, the skills of the students that remain fixed in time and that could be related to attending a school with integration. The key assumption is that these elements remain invariant during the period covered by the data panel. Nevertheless, this specification does not make it possible to correct the bias associated with the omission of each school's educational policies and practices.

[27] This methodology is implemented thanks to the high-dimension fixed effects estimator developed by Correia (2014), which makes it possible to control for time, individual, and school fixed effects simultaneously.

policy change. This is the effect of the inclusion of students with TSEN once they are diagnosed and treated and after they are allocated an additional subsidy. Theoretically one can separate the policy between (i) additional resources available at the school and (ii) diagnosis and attention to the TSEN. The extra resources could be used in not SEN specific activities and increase the academic performance of the non-SEN students. However, we aren't able to separate those effects[28].

Parameter β_2 is the estimation of the effect of PSEN inclusion conditional to the first period (2007) and $\beta_2 + \beta_4$ is the effect of PSEN inclusion conditional to the second period. β_2 can be interpreted as the effect of PSEN inclusion when these students were recognized and receive an additional subsidy but have not an attention protocol. $\beta_2 + \beta_4$ can be interpreted as the effect of PSEN when the new policy is implemented. β_4 is the effect of the policy change. This is the effect of a specific attention protocol but not an additional subsidy[29].

One concern about this specification is that allows us to exploit only the variation between students who change of schools[30]. An alternative approach could be estimate the equation (1) in a subsample of students that remains in the same school. If the results are similar, it's possible to say that the results are representative for the full sample. Another alternative equation implies replace class integration dummy variable with school-level integration dummy variables (which adopt the value 1 when there is at least one SEN student in the school)[31]. Finally, equation (1) could also be estimated by replacing the per-class integration dummy variables with the percentage of students with TSEN or PSEN in the class. This specification makes it possible to determine whether any differential effects occur depending on the number of integrated students and to evaluate if the presence of students with SEN is relevant only after a certain level[32] [33].

6. Results

[28] To separate these effects, we would need to observe a group of students treated only with additional resources (or a group of classmates treated only with new protocols of attention). Eventually, in a school j with more than one class per grade and at least with one of those classes with SEN students and other without SEN students, we would be able to observe the effect of extra resources (in the class without SEN students) and the overall effect (in the other class). However, the schools that satisfy those restrictions are few and we would not have sufficient statistical power to identify such potential effects

[29] Should be remembered that PSEN students already have extra subvention before the policy. The policy change implies attention protocols, but not extra money.

[30] Having more cohorts or extended panel data would help us smooth over random shocks and idiosyncratic differences of grades, years, or cohorts. Unfortunately, a longer panel dataset is not yet available.

[31] These specifications make it possible to address the potentially non-random allocation of students with and without SEN to the classes of each school. These equations are estimated and presented in the Results section.

[32] To find the average effect associated with one additional student, the coefficients in this expression were scaled.

[33] Alternatively, we estimate two models of Propensity Score Matching (PSM), to estimate the ATT associated to the inclusion of students with PSEN and TSEN. The first model is a simple PSM, where the variables that determine the probability of treatment are covariates at the family level (books, parents 'years of schooling, income logarithm) and school (book averages, parents' schooling, income and male / female ratio of establishment) of the year 2011. The second is originally the first, but we restrict the sample to students who in 2007 were not treated (excluding those that were treated in that year). With this, the treaties correspond to students without integration in the first period and with integration in the second period, and the controls are students without integration in any period. This allows to isolate the before / after effect. The effect of TSEN and PSEN is estimated separately. Our results indicate that the estimated "treatment" coefficients are approximately between [-1, -4] performance points. This is quite similar to the estimated OLS models and even to the FE estimator when the effect of the policy change is not considered. This indicates that PSM is not an unbiased estimator in cases like this, where precisely those not observable characteristics determine the probability of treatment. The assumption that treated and control groups are equivalent in observable and not observable would not be fulfilled. This bias is cleared by our FE estimator, which reduces the bias by unobservable constants over time. The associated tables are found in the annex section.

6.1. General results

Table 3 presents the results. Estimations for the mathematics tests show that, when controlling for time and individual fixed effects, the impact of TSEN integration during the 2007 period (β_1) is approximately -2.6 SIMCE points (0.05 SD). When simultaneously estimating time, individual, and school fixed effects, this coefficient drops to -2.1 points (0.038 SD). Integration of students with TSEN seems to have a negative effect on their classmates' mathematics performance. The effect of the policy that recognizes these students, allocates resources to them and formalizes the assistance provided ($\beta 3$) is estimated at 1.085 points (0.023 SD) in the first specification. When controlling for time, individual, and school fixed effects, the coefficient is estimated at 1.307 points (0.028 SD). In this case, we are not able to distinguish between the different components of the change of policy. We interpret this result as indicating that the policy change could revert the negative effect associated with TSEN integration. When the estimates of the second equation are considered, the effect of TSEN integration during the second period ($\beta 1 + \beta 3$) is calculated at approximately -0.7 points (0.015 SD), a statistically significant impact. In the mathematics area, the change in the treatment of these students does not fully mitigate the negative effect, but the remaining effect is of a lesser magnitude. Estimations for the reading tests show similar results. TSEN integration is observed to have a negative impact on classmates' performance; however, the effect is of a lesser magnitude than in the mathematics area. Again, the policy appears to have helped revert the negative effect associated with TSEN integration. In the reading area, the policy change with respect to TSEN students appears to fully neutralize the negative effect found.

The effect of PSEN integration (β_2) on mathematics scores during the first period it is -2 (0.04 SD). As previously established, this parameter cannot be regarded as the direct effect of PSEN integration, since these students were taught in a context tailored to their needs. Nevertheless, this result suggests that, in 2007, the policy for assisting these students was unable to fully revert the negative effect. With the same resources, but with adequate treatment, the negative effect is neutralized. With the same resources, but with adequate treatment, the negative effect is eliminated. Should be noticed that in the Reading area, the policy not only reverts a potential negative externality, but also leads to a small improvement in classmates' performance[34].

In brief, we found that once better inclusion policies are in place, the potential negative effect of being a SEN classmate student decrease or even vanish[35].

[34] To test the incidence of sorting in the main model, it is estimated using a sample that excludes courses with evidence of sorting. The model remains stable, and the estimates are in the annex, Table A.12.

[35] With regards to the estimated effects, one concern is that they are handled by a variation of the effect of the integration over time. It could be that the effects on 8th graders are smaller than the effects on 4th graders simply because negative effects from inclusion dissipate over time. In order to test this possibility, two sets of different cohort estimates were produced for level, corresponding to fourth graders (2007-2013) and eighth graders (2009-2011). Both periods cover the time period before-after the policy. Each set presents 3 models per level, which estimate the effects of the PSEN and TSEN integration before and after the policy, such as the main model. The tables A.8 and A.9, in Appendices section, present the estimates for mathematics and language. Model (1) controls for School Fixed Effects, but not for "previous performance" (the average of the previous year's grades, standardized to the school average). Model (2) does not control for School Fixed Effects but does for previous performance. Model (3) controls for School Fixed Effects and previous performance. All models include school, class and family controls, used in the main paper model. The coefficients presented in the table are analogous to those presented in the table that presents the main estimate. Although these models try to minimize identification problems, they do not account for the school choice on the part of the family, student ability, or parental concern (previous performance can be considered an approximation, although not sufficiently precise). Because of this,

Table 4 presents estimations but replacing the class integration dummy variables by the school-level integration dummy variables (which adopt the value 1 when there is at least one integrated student in the school). These specifications make it possible to address the potentially non-random allocation of students with and without SEN to the classes of each school. The results of these estimations are similar to those presented in our main model. Because of that, we can presume that the non-random allocation of students with and without SEN is not a relevant source of bias[36].

Table 5 presents the results of the estimations that employ the percentage of integrated students per class instead of dummy variables. The results for mathematics and reading are very similar.

Finally, we have performed two additional analyzes. First, to study the differentiated effects of integration by performance level. We estimate our main FE model for each performance of 2007 quartile (4 separate models). These results can be found in Annex Tables A.10. It is observed that the two quartiles with lower initial performance perceive greater punishment for TSEN and PSEN integration. However, it is also observed that this group of lower initial performance, get the highest prize associated with the policy change (TSEN * time, PSEN * time). These results are relevant for two reasons. First, they reflect that the effects of TSEN and PSEN integration primarily decrease the performance of students with lower initial performance. Second, the policy change does not benefit all groups equally, but mainly those most affected in performance by integration, showing a localized effect.

The second analysis investigates the relationship between integration and policy change. The graph on the left shows there are a punishment in performance for TSEN integration (which is what we find in our main model). In addition, students with mothers with less schooling are those who perceive greater punishment, but this negative effect is reduced as the mother is more educated, and even disappears statistically at the top. From the graph on the right the inverse is observed. While the effect of the policy has a relative prize for students of mothers with less schooling, but this is reversed as the mother's schooling increase. These results are consistent with what was found in the estimates per quartiles of initial performance and confirm that integration and policy change would have a localized effect mainly on students belonging to households with less educated fathers and mothers[37].

Finally, we perform an analysis to evaluate the possibility that our estimates are affected by the increase in resources. This is found in the Appendixes section, appendix A2.

7. Policy Implications

these results should be taken with caution. Moreover, estimated coefficients are sensitive to the model implemented. The effect of the integration by TSEN and by PSEN is higher in 4th grade than in 8th grade, in both subjects. However, the magnitude of the difference of the coefficients (4th-8th) and the statistical significance of that difference varies according to each model. Nevertheless, in all the models, the before-after relationship observed in our initial estimates is consistent, in that there is a positive effect of the policy following treatment. Still, given the identification problems of the models, it is not possible to ensure that the effects associated with the integration differ depending on grade level. However, the results obtained in our main estimate should be analyzed with caution, since it is not ruled out that the fall in the integration effect is at least partially explained by a decrease in the effect of integration as students grow.

36 This is in line with the analyses of the sorting of students with and without SEN presented in section 5.1 (Identification problems), which revealed that non-random allocation (by integration) is not a characteristic practice of the system.

37 The available data do not allow us to investigate further the mechanisms behind this result, leaving the question open for further research.

The integration of students with SEN in regular classrooms is an issue that still raises many questions, particularly in developing countries settings. The current available literature is mainly focus in developed countries and is not clear that the results could be extrapolate to different context. This study provides new evidence regarding the effect of integration on the academic performance of classmates without SEN in a developing country.

Our results show that the inclusion of PSEN and TSEN students has a small negative effect on the academic achievement of their classmates without SEN before the policy. This results are similar to Fletcher (2009) or Kristoffersen et al. (2015)[38]. When the policy change begins to operate, the negative effect of PSEN and TSEN integration on standardized test scores are fully neutralized. The additional resources and recognition seems to explain the difference with the literature that founds negative effects. This results are, on the other hand, consistent with the literature that founds no effect in developed countries context. For instance, Ruijs (2017), Friesen et al. (2010) and Aizer (2008) in the Netherlands, Canada and US context respectively. In the first two cases, the schools receive substantial additional funding for educating SEN students. In Aizer (2008), she found negative effects when the ADD students are undiagnosed, but when they are diagnosed, no negative effects are found. She also show evidence that extra resources could help to overcome negative peer effects. The results are also compatibles with the findings in Hanushek (2002), in the sense of a positive effect related with more resources for the integrated students that might overcome the negative effects.

Therefore, our paper seems to reconcile the mixed findings in the current literature. In one hand, it appears to be a small negative effect if the SEN students are not treated. In other hand, this small negative effect might be cancelled by the extra resources and accurate treatment given to SEN students.

With respect to public policy, the results presented in this article suggest that allocating additional resources for assisting integrated students with SEN is not enough to guarantee an educational context that will not harm the academic performance of these students' classmates. It is also necessary to provide assistance in the classroom in a way that considers the interplay between the student's SEN and his/her educational context. This result is remarkable because present additional elements to the previous findings in Ruijs (2017), that discusses just the effect of additional resources. This evidence should inform public policies in developing countries that have yet to generate plans for recognizing and assisting students with SEN or are only starting to do so.

8. Discussion

Our paper present new evidence of a small negative effect associate with SEN students in regular classroom on the academic achievement of their non-SEN classmates. This effect is present in the absence of policies that treat the SEN students. When this students are recognized and more resources are assigned to them the effect is dissipated. The case of PSEN illustrates that just more resources are not enough to mitigate the effect, and that an accurate treatment is also needed to do it.

[38] They investigate the case of disruptive peers with diagnosed behavioral or emotional problems, but not exactly SEN students.

Two elements must be highlighted. First, the policy change is effective since it identifies a segment of the population that was not being diagnosed (TSEN) and allocates resources and assistance to it. This improvement fully eliminates the negative effect associated with the inclusion of these students. Second, the policy change introduces a new approach to the use of resources for students who were already benefiting from the policy (PSEN). Without increasing per-student funding, these modifications fully suppress the negative effect associated with the integration of these students. This can be regarded as an increase in efficiency.

Should be considered that the findings presented are associated with the Chilean context and could be extrapolate to other developing middle-income countries (and developed countries, in the light of the literature previous findings). The policy implies that schools should invest in accurate professionals, that the country have a health system able to make a correct diagnosis of SEN students and also, that the country have a supervision system able to avoid corruption and money misuse of the school.

Moreover this study sheds light on inclusive education, it has a number of limitations. Due to insufficient data, we did not examine the specific effect of PSEN and TSEN inclusion, nor did we explore the impact of inclusion on the performance (in academic or social terms) of integrated students. In the same line, we are not able to explore the heterogeneity within both types of SEN (for instance, the difference between blind and Asperger students in the case of permanent needs).

Our work also leaves out the effect of the policy change on these students' performance. It is still essential to identify the most efficient and effective interventions for implementing integration. Questions such as where to invest, what kinds of teacher training programs are required, what materials are needed, what infrastructure would enrich integration, and other inputs are needed, constitute pending challenges for future research.

Finally, a more comprehensive study should also analyze the effect of integration on SEN students for a developing country. In the Chilean case we are not able to evaluate the academic or non-cognitive outcome of this students. It could be interesting the dynamics of the effect between type of SEN and effect on SEN and non-SEN students.

Tables

Table 1: Descriptive statistics

	Cross-section		Panel	
	2007	2011	2007	2011
With integration	0.40	0.35	0.40	0.35
	(0.49)	(0.48)	(0.49)	(0.48)
Integration of students with TSEN	0.32	0.25	0.31	0.25
	(0.47)	(0.43)	(0.46)	(0.43)
Integration of students with PSEN	0.15	0.18	0.15	0.19
	(0.36)	(0.39)	(0.36)	(0.39)
Voucher school	0.50	0.54	0.50	0.53
	(0.50)	(0.50)	(0.50)	(0.50)
Mathematics score	246	257	253	260
	(54.3)	(46.55)	(51.98)	(46.71)
Reading score	255	254	261	257
	(51.9)	(48.5)	(49.72)	(48.06)
N	156,264	172,488	116,264	116,264

Own work based on SIMCE and MINEDUC data. The table presents descriptive statistical values (means) for the variables of interest based on the initial cross-section and panel samples. The samples only comprise students without SEN. "Integration" means that students have at least one classmate with the specified SEN. For the panel, the difference in N with the cross-section sample results from the loss of observations when matching the two periods. Standard deviations in brackets.

Table 2: Descriptive statistics for Panel, differentiated by integration

	2007		2011	
	Without integration	With integration	Without integration	With integration
Reading score	265	255	261	251
	(49)	(50)	(48)	(47)
Mathematics score	258	246	265	250
	(52)	(52)	(47)	(44)
Father's schooling	12	11	12	10
	(3.33)	(3.19)	(3.35)	(3.26)
Mother's schooling	12	11	12	10
	(3.27)	(3.14)	(3.27)	(3.25)
Voucher school	0.59	0.38	0.66	0.31
	(0.49)	(0.49)	(0.48)	(0.46)
Gender (male=1)	0.48	0.49	0.48	0.48
	(0.5)	(0.5)	(0.5)	(0.5)
Income (dollars)	620	470	775	518
	(571)	(429)	(704)	(486)
N	70,068	46,196	75,144	41,120

Own work based on SIMCE and MINEDUC data. The table presents descriptive statistical values (means) for the variables of interest of the 2007-2011 panel sample, differentiated by integration. The samples only comprise students without SEN. For the purposes of this table, integration means that students have at least one classmate with either TSEN or PSEN. Standard deviations in brackets. Income in 2010 dollars: 1 dollar=510 pesos.

Table 3: Panel with fixed effects

	Mathematics		Reading	
TSEN*Time (β_3)	1.085***	1.307***	0.769**	0.716**
	(0.313)	(0.331)	(0.321)	(0.340)
PSEN*Time (β_4)	1.892***	1.952***	2.021***	2.120***
	(0.387)	(0.410)	(0.397)	(0.420)
Class with TSEN integration (β_1)	-2.591***	-2.067***	-1.477***	-0.581*
	(0.266)	(0.298)	(0.273)	(0.305)
Class with PSEN integration (β_2)	-1.704***	-2.030***	-1.564***	-1.429***
	(0.338)	(0.370)	(0.348)	(0.383)
$\beta_1 + \beta_3$	-1.506***	-0.760**	-0.708**	0.135
	(0.292)	(0.328)	(0.297)	(0.331)
$\beta_2 + \beta_4$	0.188	-0.078	0.457	0.694*
	(0.312)	(0.34)	(0.32)	(0.349)
Individual and time fixed effects (FE)	Yes	Yes	Yes	Yes
School FE	No	Yes	No	Yes
Family characteristics	Yes	Yes	Yes	Yes
Class and school characteristics	Yes	Yes	Yes	Yes
N	232,527	231,218	232,527	231,218

This table presents estimations for the 2007-2011 panel. Columns 1 and 2 use mathematics SIMCE scores as the dependent variable. Columns 3 and 4 use reading SIMCE scores as the dependent variable. Household characteristics include controls that vary over time, such as books, income, and parents' schooling. School characteristics include controls by type (municipal, voucher, or private). Class characteristics comprise controls by parents' average schooling, class size, and evidence of SEN sorting. Standard errors calculated per individual cluster over time. * p<0.1, ** p<0.05, *** p<0.01.

Table 4: Panel with fixed effects – School-level integration

	Mathematics		Reading	
School TSEN*Time (β_3)	0.41	0.49	-0.17	-0.39
	(0.30)	(0.32)	(0.30)	(0.32)
School PSEN*Time(β_4)	1.20***	1.43***	1.68***	1.92***
	(0.33)	(0.35)	(0.34)	(0.36)
School TSEN (β_1)	-1.76***	-1.26***	-0.59**	0.39
	(0.25)	(0.29)	(0.26)	(0.30)
School PSEN (β_2)	-1.18***	-1.63***	-1.68***	-1.67***
	(0.29)	(0.33)	(0.29)	(0.34)
$\beta_1 + \beta_3$	-1.35***	-0.77**	-0.76***	0.00
	(0.27)	(0.30)	(0.28)	(0.31)
$\beta_2 + \beta_4$	0.02	-0.19	0.01	0.25
	(0.31)	(0.34)	(0.32)	(0.35)
Individual and time fixed effects (FE)	Yes	Yes	Yes	Yes
School FE	No	Yes	No	Yes
Family characteristics	Yes	Yes	Yes	Yes
Class and school characteristics	Yes	Yes	Yes	Yes
N	232,527	231,218	232,527	231,218

This table presents estimations for the 2007-2011 panel. Columns 1 and 2 use mathematics SIMCE scores as the dependent variable. Integration dummy variables (School SEN) take value 1 when the school enrolls at least one student with SEN and 0 if it does not. Columns 3 and 4 use reading SIMCE scores as the dependent variable. Household characteristics include controls that vary over time, such as books, income, and parents' schooling. School characteristics include controls by type (municipal, voucher, or private). Class characteristics comprise controls by parents' average schooling, class size, and evidence of SEN sorting. Standard errors calculated per individual cluster over time. * p<0.1, ** p<0.05, *** p<0.01.

Table 5: Panel with fixed effects – Share SEN as regressor

	Mathematics		Reading	
Share TSEN*Time (β_3)	0.83***	0.93***	0.56***	0.43*
	(0.21)	(0.23)	(0.21)	(0.23)
Share PSEN*Time (β_4)	3.02***	3.47***	2.93***	3.23***
	(0.38)	(0.43)	(0.37)	(0.43)
Share TSEN (β_1)	-1.73***	-1.17***	-1.18***	-0.66**
	(0.22)	(0.28)	(0.22)	(0.28)
Share PSEN (β_2)	-1.94***	-2.33***	-1.66***	-1.48***
	(0.34)	(0.43)	(0.34)	(0.44)
$\beta_1 + \beta_3$	-0.90***	-0.24	-0.62***	-0.23
	(0.21)	(0.25)	(0.22)	(0.25)
$\beta_2 + \beta_4$	1.09***	1.13***	1.27***	1.75***
	(0.36)	(0.43)	(0.36)	(0.43)
Individual and time fixed effects (FE)	Yes	Yes	Yes	Yes
School FE	No	Yes	No	Yes
Family characteristics	Yes	Yes	Yes	Yes
Class and school characteristics	Yes	Yes	Yes	Yes
N	232,527	231,218	232,527	231,218

This table presents estimations for the 2007-2011 panel. The coefficients presented are evaluated for the effect of one additional student with PSEN or TSEN in a 15-student class (the average size of the classes in the sample used in the estimation). Columns 1 and 2 use mathematics SIMCE scores as the dependent variable. Columns 3 and 4 use reading SIMCE scores as the dependent variable. Household characteristics include controls that vary over time, such as books, income, and parents' schooling. School characteristics include controls by type (municipal, voucher, or private). Class characteristics comprise controls by parents' average schooling, class size, and evidence of SEN sorting. Standard errors calculated per individual cluster over time. * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Robustness analysis

	Mathematics	Reading
TSEN*Time	2.496***	1.685**
	(0.501)	(0.509)
PSEN*Time	2.823***	3.029***
	(0.533)	(0.544)
Individual and time fixed effects (FE)	Yes	Yes
Family characteristics	Yes	Yes
Class and school characteristics	Yes	Yes
Sample	SEN in both periods	SEN in both periods
N	80,994	80,994

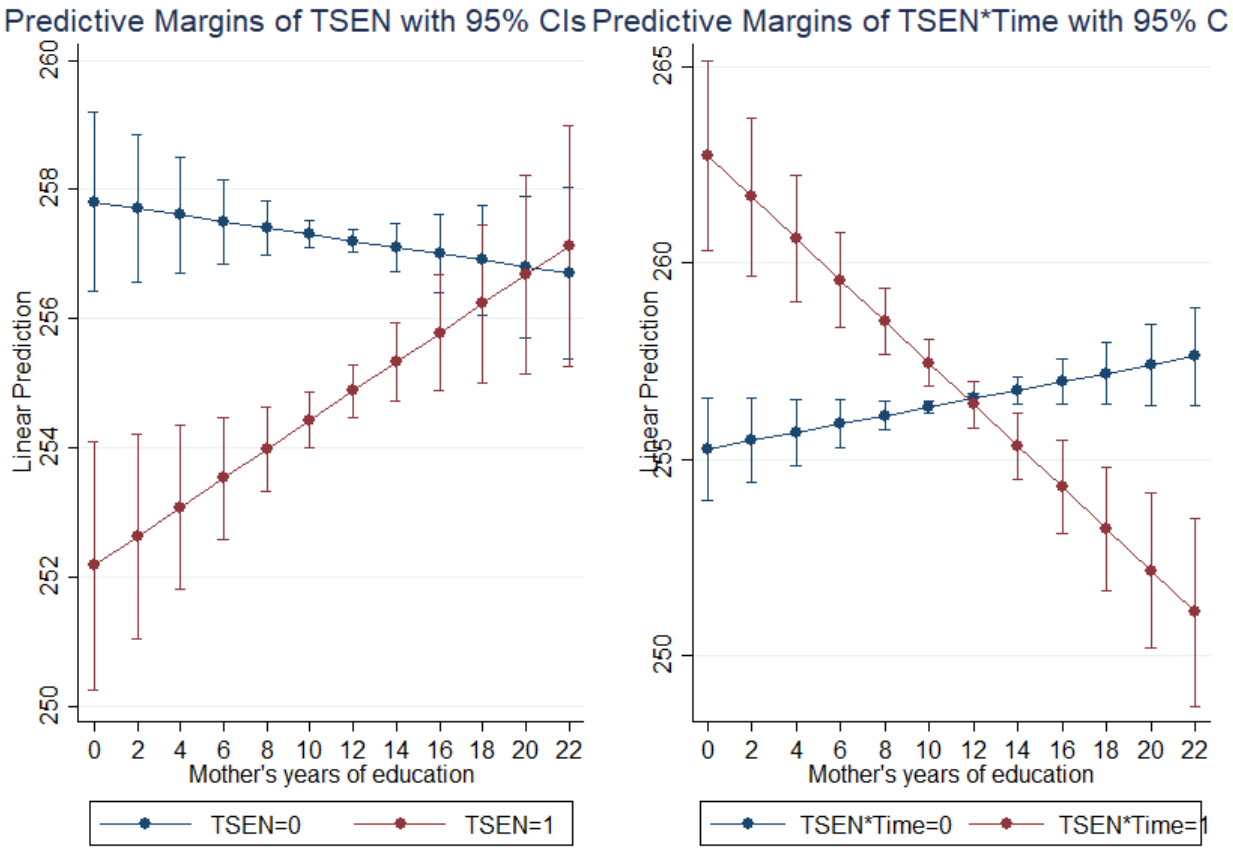
This table presents estimations for the 2007-2011 panel. The sample was limited to schools that enrolled students with SEN in both periods. The treatment group is still comprised by students who had a classmate with SEN. The control group is composed of students who did not have a classmate with SEN in schools that served students with SEN. Thus, the difference between these groups should only be due to the new treatment for students with PSEN and TSEN. Standard errors calculated per individual cluster over time. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 7: Effect on grade retention rate

Dependent variable	Retention rate
TSEN*Time (β_3)	0.006
	(0.0023)
PSEN*Time (β_4)	-0.0122***
	(0.0025)
School with TSEN integration (β_1)	0.0006
	(0.0022)
School with PSEN integration (β_2)	0.0114***
	(0.0025)
$\beta_1 + \beta_3$	0.0012
	(0.0022)
$\beta_2 + \beta_4$	-0.0008
	(0.0023)
Time FE	Yes
School FE	Yes
N	8,810

This table presents estimations for the 2007-2011 panel at the school level. The following school-level characteristics are controlled for: father's and mother's educational level, household income, and books in the home. Standard errors calculated per school cluster over time. * p<0.1, ** p<0.05, *** p<0.01.

Graph 1:



In the left panel of the graph 1, we present the linear prediction of the model 1 evaluating TSEN*Mother's years of education interaction. Predictions for TSEN = 0 and TSEN = 1 are presented. The right panel is analogous but evaluating the policy change (TSEN * time) on the different levels of Mother's years of education.

References

Aizer, A. (2008). Peer effects and human capital acquisition: The externalities of ADD. NBER Working Paper 14354, National Bureau of Economic Research.

Angrist, J., & Lang, K. (2004). Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. The American Economic Review, 94(5), 1613-1634. Retrieved from http://www.jstor.org/stable/3592836

Clotfelter, C., Ladd, H., & Vigdor, J. (2006). Teacher-Student Matching and the Assessment of Teacher Effectiveness. Journal of Human Resources, 41, 778-820.

Contreras, D., Sepúlveda, P., & Bustos, S. (2010). When schools are the ones that choose: The effects of screening in Chile. Social Science Quarterly, 91(5), 1349-1368.

Correia, S. (2016). REGHDFE: Stata module to perform linear or instrumental-variable regression absorbing any number of high-dimensional fixed effects.

Fletcher, J. M. (2009). The effects of inclusion on classmates of students with special needs: The case of serious emotional problems. Education Finance and Policy, 4(3), 278-299.

Friesen, J., Hickey, R., & Krauth, B. (2010). Disabled peers and academic achievement. Education Finance and Policy, 5(3), 317-348

Gottfried, M. A. (2014). Classmates with disabilities and students' noncognitive outcomes. Educational Evaluation and Policy Analysis, 36(1), 20-43.

Gottfried, Michael & Kirksey, J. & Wright, Adam. (2019). Same-Race Student–Teacher: Comparing Outcomes for Kindergartners With and Without Disabilities. Remedial and Special Education. 40. 225-235. 10.1177/0741932518810870.

Gottfried, Michael & Kirksey, J.. (2019). General education teachers' math instructional practices in kindergarten classrooms with and without students with emotional and behavioral disabilities. Teaching and Teacher Education. 77. 309-320. 10.1016/j.tate.2018.10.019.

Gould, E. D., Lavy, V., & Daniele Paserman, M. (2009). Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence. The Economic Journal, 119(540), 1243-1269.

Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2002). Inferring program effects for special populations: Does special education raise achievement for students with disabilities? Review of Economics and Statistics, 84(4), 584-599.

Houtenville, Andrew & Conway, Karen. (2008). Parental Effort, School Resources, and Student Achievement. Journal of Human Resources. 43. 437-453. 10.2307/40057353.

Jordan, Anne & Schwartz, Eileen & McGhie-Richmond, Donna. (2009). Preparing teachers for inclusive classrooms. Teaching and Teacher Education. 25. 535-542. 10.1016/j.tate.2009.02.010.

Keslair, F., Maurin, E., & McNally, S. (2012). Every child matters? An evaluation of "Special Educational Needs" programmes in England. Economics of education review, 31(6), 932-948.

Kristoffersen, J. H. G., Krægpøth, M. V., Nielsen, H. S., & Simonsen, M. (2015). Disruptive school peers and student outcomes. Economics of Education Review, 45, 1-13.

Raitano, M., & Vona, F. (2013). Peer heterogeneity, school tracking and students' performances: Evidence from PISA 2006. Applied Economics, 45(32), 4516-4532.

Ruijs, Nienke & Peetsma, Thea. (2009). Effects of inclusion on students with and without special educational needs reviewed. Educational Research Review. 4. 67-79. 10.1016/j.edurev.2009.02.002.

Ruijs, N. (2017). The impact of special needs students on classmate performance. Economics of Education Review, 58, 15-31.

Sacerdote, B. (2012). When the saints go marching out: Long-term outcomes for student evacuees from Hurricanes Katrina and Rita. American Economic Journal: Applied Economics, 4(1), 109-35.

Salend, S. J., & Garrick Duhaney, L. M. (1999). The impact of inclusion on students with and without disabilities and their educators. Remedial and special education, 20(2), 114-126.

Schwartz, A. E., Hopkins, B. G., & Stiefel, L. (2019). The effects of special education on the academic performance of students with learning disabilities (EdWorkingPaper: 19-86). Annenberg Institute at Brown University.

Treviño, E., Valenzuela, J. P., & Villalobos, C. (2014). Segregación académica y socioeconómica al interior de la escuela. Análisis de su magnitud, evolución y principales factores explicativos. Santiago: CPCE-CIAE.

Treviño, E., Valenzuela, J. P., & Villalobos, C. (2016). Within-school segregation in the Chilean school system: What factors explain it? How efficient is this practice for fostering student achievement and equity?. Learning and Individual Differences, 51, 367-375.

Appendixes

Appendix A.1: Detailed identification problems

A1.1 Parental school selection

The first identification problem is the parental school selection. Parents can choose schools with or without an inclusive educational project[39]. Public schools (which are free) offer educational projects with and without inclusion. If parents' chosen educational approach is systematically related to other relevant but unobserved variables, such as their participation in students' education, leaving out this variable could generate bias in the estimation of the parameter of interest. Longitudinal data are used to solve this problem. An individual-level fixed effects estimation is implemented, which partially control for parental participation (or other unobserved variables, such as the cultural capital of the household). The key assumption is that those are variables that remains stable throughout the period studied. Since the horizon of the panel is only four years, the assumptions made are plausible for the sample[40].

There is evidence that, when the amount of resources allocated to a school increases, parental effort decreases (Houtenville & Convay, 2008). This crowding out effect is not clearly present in the Chilean model. Indeed, it is not evident that parents of students without SEN visualize the increase in resources, in this case linked to the recognition of TSEN students. It should be noted that, in the Chilean system, schools are mainly funded through the regular per-student voucher provided by the authorities, parental copayment (when demanded by the school) and other funding provided by the municipality (in municipal schools) or the private owner (in voucher schools).

A1.2 School preferences

A second identification problem is the omission of school preferences. The school's choice of developing an integration project (SIP) is likely to be strongly aligned with (i) their efforts, policies and practices aimed at serving a more diverse students; and (ii) the type of student that they enroll through their selection process[41]. If these parallel practices (different from those specifically associated with assistance for SEN students and the development of a SIP) impact academic performance and are correlated with the decision of integration, the results may be biased. As with the previous identification problem, the use of a fixed-effects model (this time at the school level) allow us to eliminate components associated with each school which are invariant over time. Again, the assumption behind this adjustment is that the period studied is short enough to guarantee that the school's policies, practices, and selection methods will not undergo any significant

[39] Parents also decide the school to which they send their children, which can be voucher or municipal, for-profit or non-profit, selective or non-selective (with various types of selection criteria) and require copayment (in voucher schools). These and other elements determine the makeup of each school.
[40] This period is likely to be too short to reveal significant changes in parents' behavior.
[41] For instance, schools that do not develop a SIP are not only choosing not to admit students with SEN: they are also able to select students with a better family environment.

changes. Using this approach, we are able to exploit only the variation of those students who move between schools[42].

A1.3 SEN Students Sorting

The potential placement of students with SEN in specific classes in their schools could be another potential identification problem. If teachers and administrative teams devote special efforts to these classes and develop special methodologies for them, bias may be introduced into the estimation of the parameter of interest. To rule out the presence of this pattern, a type of sorting analysis was implemented based on the strategy initially used by Clotfelter, Ladd, and Vigdor (2006) and later by Trevino et al. (2014). These authors test sorting by socioeconomic characteristics and performance. We applied this approach in order to test sorting by percentage of integrated students[43]. Our results (presented in table A.5, Appendix section) show that, in fourth grade (2007), 2.2% of the total number of classes in the panel sample display evidence of sorting by SEN. 0.4% of the classes display evidence of sorting by PSEN, while 1.5% show sorting by TSEN. This means that 2.2% of students in the panel sample are in classes where there is evidence of sorting by SEN, while 0.6% and 2% are in classes with evidence of sorting by PSEN and TSEN respectively. For eighth grade (2011), 3% of the total number of classes in the sample display evidence of sorting by SEN. Classes with sorting by PSEN amount to 1% of the total number of classes, while those with sorting by TSEN represent 2%. This means that 3.2% of students are in classes with evidence of sorting by SEN: 1.1% are in classes with sorting by PSEN and 2.2% are in classes with sorting by TSEN. Although there does not seem to be a systematic attempt to place students with SEN in specific classes, we include a variable that indicates if the null hypothesis is rejected in that class[4445].

A1.4 Non-Random Assignment

Lastly, an identification problem that must be addressed is that students without SEN could be assigned non-randomly to classes with SEN. For instance, students could be placed in a class according to their prior academic performance or socioeconomic characteristics. A non-random distribution could affect student achievement not only due to peer characteristics (which can be mitigated through class-level controls), but also because teachers could perform differently in each

[42] An additional exercise of robustness may be estimate only on the subsample of students who remain in the same school. If these results are similar, we can assume that the estimations are valid for all students.

[43] It consists of a statistical test (chi2) of the null hypothesis that the percentage of students with SEN in a school's classes (in schools with two or more classes) is the same in all classes. If this test is not significant at the 10% level, that is, if a class has a much larger share of students with SEN than the rest of the school, there is a reason to believe that students with SEN are being sorted.

[44] The data could be sorted by SEN type –both disability type and level of need. Unfortunately, databases include neither the "level" of a student's disability nor the specific disability affecting him/her.

[45] Our analysis of sorting by SEN stands in contrast to the evidence of academic sorting documented in the Chilean educational system (see Trevino et. al. 2014, 2016). In 2009, 30% of students in the whole educational system attended schools where academic sorting was implemented (a value that only reached 13% in OECD countries). We believe that the large difference with the 3.2% of students who are in classes with evidence of sorting by SEN can be explained by two factors. First, there is a limit on the number of students with SEN that schools can place in a single class. Second, we hypothesize that the optimal behavior that schools have identified is to distribute students with SEN equally among classes. A larger concentration of students with SEN results in increased assistance requirements. This situation can entail an excessive burden on teachers, potentially causing a negative impact on academic performance.

class. To check for non-randomness, we estimated a linear probability model (LPM) of the likelihood of being in a class that integrates students with SEN. If the likelihood that a student will join a class that integrates SEN students depends on his/her characteristics and past performance, selection can be said to be non-random. The equation to be estimated is the following[46]:

$$Class\ SEN_{ij} = \alpha + \gamma * X_{ij} + \lambda_j + \varepsilon_{ij}$$

Where $Class\_SEN_{ij}$ adopts the value 1 when student "i" in school "j" is in a class with at least one classmate with SEN. X_{ij} is the subject's characteristics vector and his/her past performance. Lastly, λ_j is the coefficient associated with school "j".

Table A.6 presents the results of these estimations for the years 2007 and 2011. For both years, the likelihood of being in a class with TSEN classmates is not significantly explained by any of the variables. For the rest of the models, in 2007, the only variables found to be relevant are the number of books and grade retention[47]. The evidence found suggests that the placement of students without SEN in the same classes as students with SEN behaves in a manner that resembles random allocation. Therefore, this cannot be said to be a relevant identification problem[48]. Nevertheless, these results must be analyzed cautiously given what has been pointed out beforehand.

[46] The controls used were student characteristics (age, gender), socioeconomic status (income, number of books in the household, years of schooling of each parent), and prior academic performance (average grade in all previous levels, standardized on a per-school basis; along with a dummy variable that indicates whether the student has repeated a grade). The model also adds a control for each school, which enables us to study the conditional probability of attending a specific school.

[47] One additional book reduces the student's likelihood to be in an integrated class, but the effect is vanishingly small and only significant for PSEN integration. If the student has repeated a grade before 2007, he/she is 1% more likely to be in a class that integrates according to SEN or PSEN status. As for 2011, academic achievement in the previous year only has an impact on the likelihood of being in a class with PSEN integration. The effect is positive and significant, but very small.

[48] We also estimated a model identical to the one presented, but using the ratio *integrated students/total number of students in class* as the dependent variable. The conclusions are similar to the model presented.

Appendix A.2: Robustness analysis

The results presented include the effect of the policy change, that is, the increase in resources and the improved assistance protocols (with both policy approaches benefiting TSEN students and only the latter benefiting PSEN students). A potential concern about these results is that the effect of the policy change may be caused by the increase in resources and not by the better treatment or the increase of SEN-specific resources; that is, schools may introduce improvements that benefit not only students with SEN, but also the rest of their peers.

If we assume that the additional funding for a school that serves students with SEN (SEN school) has an equivalent effect in classes with and without SEN integration (SEN and non-SEN class respectively), then the only difference between a SEN class and a non-SEN class of a SEN school would be the new protocols for SEN students. Therefore, to find the effect of the new protocols, we can calculate the difference between SEN and non-SEN classes conditional on being in a SEN school. On the contrary, if the assumption that the extra funds have an equivalent effect on both groups of classes is not satisfied (because the impact is greater in those with SEN students, where the money is spent on SEN-specific activities), then the effect of the policy must be attributed to (i) the protocol and (ii) the increased resources for treating students with SEN.

To test the above scenario, we estimate restricting the sample to students who attend SEN school during both periods. This means that (i) the schools will be similar regarding their likelihood to integrate and with respect to non-observables that influence their choice, and (ii) that parents, conditional to selecting an integrated school, will be equal in terms of non-observables, and (iii) that a student will randomly be assigned to a class with (or without) SEN students. Therefore, if differences exist within this group, they can only be due to improvements in treatment and service delivery protocols.

Results are presented in table 6. It is observed that the coefficients are positive on average, which indicates that a positive difference exists between these groups. Then, since the groups are equivalent on average except for the specific treatment of students with SEN, this difference can be specifically attributed to the effect of the new treatment protocol. This suggests that the results yielded by our main estimation are not only due to the additional funding, that there is an effect connected to the policy change specific to SEN activities.

In order to reduce the identification problems associated with the increase in resources, we have also estimated a set of models that limit the variation between both periods, of the proportion of SEN students (TSEN / PSEN).

For this analysis, we keep in the sample only schools that are present in the databases of both periods. Then, we calculate the variations between both periods, of the proportion of students with TSEN, PSEN or all SEN (3 different variations are calculated). We estimate separately for TSEN, PSEN and SEN, our main FE model, but we restrict each one to 3 cases: sample with students from schools where the variation was less than 10%, where it was less than 5%, and where it was less than 3% (the last level where statistical significance of the parameters of interest is found).

The first thing to note is that the variation in the proportion over time, at the school level, is reduced (1.4% on average for TSEN, almost 1% for PSEN and 2% for SEN). This is due to policy restrictions regarding the number of students integrated. Consequently, the restrictions imposed do not exclude as much sample as might be expected. The different models show that the effects are reduced while the required variation is smaller, but the relationship observed in the initial model is maintained, where the policy change cancels the effect of inclusion. This is a good result of the robustness of the model, as it would indicate that our estimates would not be handled by variation in the time of TSEN / PSEN students in schools. Tables with the models estimated in the annex: Tables A.11.

Table A.1 Comparison of TSEN identification approaches

Type of identification		TSEN students according to 2013 database	
		No	Yes
TSEN students in 2011 according to our methodology	No	-	2753 (a)
	Yes	2557 (b)	1980 (c)

Source: Own work based on administrative data. The table compares the two approaches to the identification of students with TSEN. The rows present the identification according to the methodology discussed in the paper (students identified as having SEN but not PSEN in 2011). The columns present an alternative identification (those with TSEN in 2013 are also identified as having TSEN in 2011). Each cell was defined using a letter for easier reference. (a) displays the matches between both approaches. (b) shows the cases that we identified, and which were not identified by the 2013 database. (c) shows the students who have TSEN according to the 2013 database, but who we did not identify as such.

Table A.2 Distribution of students with SEN

Students	Cross-section data					
	2007			2011		
	Municipal	Voucher	Total	Municipal	Voucher	Total
	(1)	(2)	(3)	(4)	(5)	(6)
Total	80,794	80,851	161,645	84,266	95,073	179,339
With SEN	3,655	1,726	5,381	4,734	2,134	6,868
	(4.52%)	(2.13%)	(3.33%)	(5.62%)	(2.24%)	(3.83%)
With PSEN	1,122	345	1,467	1,835	516	2,351
	(1.4%)	(0.4%)	(0.9%)	(2.2%)	(0.5%)	(1.3%)
With TSEN	2,533	1,381	3,914	2,899	1,618	4,517
	(3.1%)	(1.7%)	(2.4%)	(3.4%)	(1.7%)	(2.5%)

Source: Own work based on administrative data. The table shows the distribution of students with SEN by school type and year. The initial cross-section samples are used. Columns (3) and (6) show the total number of students for each year. The percentage relative to the total number of students for that year and for that school type is shown in brackets.

Table A.3 Distribution of students with and without SEN.

Type of student	Cross-section data					
	2007			2011		
	Municipal	Voucher	Total	Municipal	Voucher	Total
Panel A: Students without SEN						
Total without SEN	77,139	79,125	156,264	79,532	92,939	172,471
	(49%)	(51%)		(46%)	(54%)	
In both periods	59,682	60,982	120,664	56,238	64,426	120,664
	(49%)	(51%)		(47%)	(53%)	
Panel B: Students with SEN						
Total with SEN	3,655	1,726	5,381	4,734	2,134	6,868
	(68%)	(32%)		(69%)	(31%)	
In both periods	2,072	978	3,050	2,028	1,022	3,050
	(68%)	(32%)		(67%)	(33%)	
Panel C: Students with PSEN						
Total with PSEN	1,122	345	1,467	1,835	516	2,351
	(77%)	(23%)		(78%)	(22%)	
In both periods	614	171	785	610	175	785
	(78%)	(22%)		(78%)	(22%)	
Panel D: Students with TSEN						
Total with TSEN	2,533	1,381	3,914	2,899	1,618	4,517
	(65%)	(35%)		(64%)	(36%)	
In both periods	1,458	807	2,265	1,418	847	2,265
	(64%)	(36%)		(63%)	(37%)	

Source: Own work based on administrative data. The table shows the frequency of students by school type and year. It is divided into four panels: students with SEN, students without SEN, students with PSEN, and students with TSEN. Each panel presents the total number of students (row 1) and the number of students who meet the condition of appearing in both periods in that integration category ("In both periods" - row 3). Percentages relative to the total of each category in that period are shown in brackets.

Table A.4 Statistics for classes with integrated students

	Panel Data	
Statistical value	2007	2011
Panel A: Total number of classes		
Total number of classes in the system	7,911	7,688
% of classes with SEN integration	38.2%	39.6%
% of classes with PSEN integration	15.0%	21.2%
% of classes with TSEN integration	29.5%	28.3%
Panel B: Class sizes		
Class size without SEN	14.32	16.17
Class size with SEN	15.30	13.50
Class size with PSEN	14.64	13.43
Class size with TSEN	15.68	13.39
Panel C: Total number of students with SEN		
Number of students with SEN	1.75	2.17
Number of students with PSEN	1.21	1.38
Number of students with TSEN	1.65	2.01

Source: Own work based on administrative data. The table displays statistical data about classes and their integration status. An integrated class was defined as one with at least one SEN student. For instance, a class with PSEN integration is one with at least one PSEN student. Panel A shows the total number of classes and the distribution of those with SEN, PSEN, or TSEN integration. Panel B displays class sizes relative to integration. Finally, panel C displays the average number of students with SEN (or TSEN or PSEN) in an integrated class.

Table A.5 Statistical evidence of sorting by SEN

	Panel Data	
Statistical value	2007	2011
% of classes that display evidence of sorting by:		
SEN	2.2%	3.0%
PSEN	0.4%	1.0%
TSEN	1.5%	2.0%
% of students in a class that displays evidence o sorting by:		
SEN	2.2%	3.2%
PSEN	0.6%	1.1%
TSEN	2.0%	2.2%

Source: Own work based on administrative data. The table presents the results of the test used to detect the sorting of students with SEN within schools. It consists of a statistical test (chi2) of the null hypothesis that the percentage of students with SEN in a school's classes (in schools with two or more classes) is the same in all classes. If this test is not significant at the 10% level, that is, if a class has a much larger or smaller share of students with SEN than the rest of the school, there is reason to believe that students with SEN are being sorted. The upper panel displays the percentage of classes where there is evidence of sorting, relative to the total number of classes in the sample. The lower panel displays the percentage of students in classes where there is evidence of sorting, relative to the total number of classes in the sample. The sample comprises the panel data used in the main estimation.

Table A.6 Linear probability models of the likelihood of attending an integrated class

	2007			2011		
	SEN	PSEN	TSEN	SEN	PSEN	TSEN
	(1)	(2)	(3)	(4)	(5)	(6)
Age	0.00044	0.00137	0.000948	-0.000127	0.0000267	-0.000135
	(0.0013)	(0.00150)	(0.00148)	(0.000109)	(0.0000938)	(0.000101)
Gender (1=male)	0.0018	0.00157	0.00138	-0.00172	-0.000747	-0.00158
	(0.0015)	(0.00120)	(0.00145)	(0.00133)	(0.00113)	(0.00126)
Father's schooling	0.0000669	0.000116	-0.000228	-0.000192	-0.0000698	-0.000340
	(0.000373)	(0.000281)	(0.000342)	(0.000287)	(0.000255)	(0.000263)
Mother's schooling	-0.000248	0.0000204	-0.000308	-0.000160	-0.0000616	-0.0000228
	(0.000363)	(0.000290)	(0.000338)	(0.000290)	(0.000264)	(0.000270)
Income logarithm	-0.00216	0.000230	-0.00201	-0.00203	-0.000734	-0.00120
	(0.00144)	(0.00114)	(0.00135)	(0.00127)	(0.00110)	(0.00117)
Grade retention	0.0105**	0.00972**	0.00521	0.00455	0.00273	0.00538
	(0.00436)	(0.00400)	(0.00424)	(0.00444)	(0.00443)	(0.00416)
Books in the household	-0.0000154	-0.0000810**	0.0000200	-0.00000734	0.0000102	-0.0000185
	(0.0000366)	(0.0000317)	(0.0000272)	(0.0000215)	(0.0000184)	(0.0000196)
Performance in 2005	0.00176	0.000539	0.000525	-0.000	-0.000***	0.000
	(0.00319)	(0.00224)	(0.00309)	(0.000)	(0.000)	(0.000)
Performance in 2006	-0.00276	-0.00212	-0.00116	-0.000403	-0.00212	-0.0000501
	(0.00306)	(0.00221)	(0.00296)	(0.00208)	(0.00200)	(0.00188)
Performance in 2007				-0.00105	0.00190	-0.00193
				(0.00234)	(0.00211)	(0.00208)
Performance in 2008				0.000	-0.000	0.000
				(0.000)	(0.000)	(0.000)
Performance in 2009				0.00144	0.00197	0.00150
				(0.00226)	(0.00193)	(0.00204)
Performance 2010				-0.00229	-0.00405**	-0.000887
				(0.00194)	(0.00175)	(0.00172)
N	146410	146410	146410	146524	146524	146524

Source: Own work based on administrative data. Standard errors in brackets. The table presents the results of a regression model of the likelihood of attending an integrated class for 2007 and 2011. Columns (1) and (4) show the likelihood of attending a class with SEN integration. Columns (2) and (5) show the likelihood of attending a class with PSEN integration. Columns (3) and (6) show the likelihood of attending a class with TSEN integration. In all regressions, observable student characteristics were controlled for. The estimations in fourth grade (2007) controlled for prior academic performance in 2005 and 2006 (second and third grade respectively). The estimations in eighth grade (2011) controlled for prior academic performance in 2005, 2006, 2007, 2008, 2009, and 2010 (second, third, fourth, fifth, sixth, and seventh grade respectively). The academic performance control used was the student's final annual grade relative to the grades of his/her class.

Tables A.7 Propensity Score Matching (PSM) estimations for TSEN and PSEN integration

ATT estimation with Nearest Neighbor Matching method. Bootstrapped standard errors. Model 1 with whole sample.

	N. treat	N.contr	ATT	Std.Err.	t
TSEN on Math	29.087	21.691	-1.969	0.637	-3.093
TSEN on Reading	29.087	21.691	-1.679	0.425	-3.954
PSEN on Math	21.905	17.165	-1.951	0.239	-8.16
PSEN on Reading	21.905	17.165	-0.877	0.542	-1.619

Source: Own work based on administrative data. Covariables for matching = family level: books, parents' scholarships and income; school level: means of books, parents' scholarships, income and gender share). Sample restricted to common support.

ATT estimation with Nearest Neighbor Matching method. Bootstrapped standard errors. Model 2 with restricted sample (without any SEN in 2007 (first period)).

	N. treat	N.contr	ATT	Std.Err.	t
TSEN on Math	11.165	9.606	-4.202	0.444	-9.461
TSEN on Reading	11.165	9.606	-4.043	1.214	-3.331
PSEN on Math	13.739	11.671	-2.051	0.312	-6.576
PSEN on Reading	13.739	11.671	-0.996	0.827	-1.205

Source: Own work based on administrative data. Covariables for matching = family level: books, parents' scholarships and income; school level: means of books, parents' scholarships, income and gender share). Sample restricted to common support

Table A.8: Different cohorts for the same level (Mathematics)

	4th grade (2007-2013)			8th grade (2009-2011)		
	(1)	(2)	(3)	(1)	(2)	(3)
TSEN*Time (β_3)	4.507***	7.676***	4.607***	2.977***	2.215**	1.713*
	(0.946)	(1.013)	(0.996)	(1.071)	(1.027)	(1.000)
PSEN*Time (β_4)	5.557***	8.705***	7.499***	3.742**	2.362	2.518*
	(1.051)	(1.164)	(1.122)	(1.504)	(1.507)	(1.405)
Class with TSEN integration (β_1)	-4.423***	-3.657***	-3.436***	-4.831***	-2.059**	-1.694*
	(0.684)	(0.714)	(0.692)	(0.986)	(0.931)	(0.910)
Class with PSEN integration (β_2)	-5.458***	-5.161***	-4.556***	-3.800***	-2.250	-2.026
	(0.794)	(0.872)	(0.831)	(1.410)	(1.389)	(1.304)
Time effects	Yes	Yes	Yes	Yes	Yes	Yes
School FE	No	Yes	Yes	No	Yes	Yes
Previous performance	Yes	No	Yes	Yes	No	Yes
N	331880	340083	331880	299438	310060	299438

The table shows estimates for mathematics. The samples are composed of different cohorts for the same level. They correspond to two sets of estimates: for fourth grade (2007-2013) and for eighth grade (2009-2011). Both periods cover the time period before-after the policy. The models (1) control by School FE, but not by Previous performance (average of the previous year's grades, standardized to the school average). The models (2) do not control by School FE, but by Previous performance. Models (3) controlled by School FE and Previous performance. All models include school, class and family controls, used in the main paper model. The coefficients presented in the table are analogous to those presented in the table that presents the main estimate. Standard errors calculated by school cluster. * p<0.1, ** p<0.05, *** p<0.01.

Table A.9: Different cohorts for the same level (Reading)

	4th grade (2007-2013)			8th grade (2009-2011)		
	(1)	(2)	(3)	(1)	(2)	(3)
TSEN*Time (β_3)	3.226***	5.267***	2.417***	1.603	1.651	0.973
	(0.792)	(0.851)	(0.825)	(1.106)	(1.096)	(1.070)
PSEN*Time (β_4)	3.747***	6.023***	4.881***	2.638*	1.742	2.297
	(0.901)	(0.968)	(0.927)	(1.511)	(1.570)	(1.520)
Class with TSEN integration (β_1)	-3.389***	-2.238***	-1.988***	-3.672***	-1.901*	-1.594*
	(0.587)	(0.612)	(0.589)	(-1.004)	(0.974)	(0.960)
Class with PSEN integration (β_2)	-4.564***	-3.413***	-2.868***	-1.904	-1.560	-1.393
	(0.698)	(0.771)	(0.732)	(1.372)	(1.423)	(1.383)
Time effects	Yes	Yes	Yes	Yes	Yes	Yes
School FE	No	Yes	Yes	No	Yes	Yes
Previous performance	Yes	No	Yes	Yes	No	Yes
N	331880	340083	331880	299438	310060	299438

The table shows estimates for reading. The samples are composed of different cohorts for the same level. They correspond to two sets of estimates: for fourth grade (2007-2013) and for eighth grade (2009-2011). Both periods cover the time period before-after the policy. The models (1) control by School FE, but not by Previous performance (average of the previous year's grades, standardized to the school average). The models (2) do not control by School FE, but by Previous performance. Models (3) controlled by School FE and Previous performance. All models include school, class and family controls, used in the main paper model. The coefficients presented in the table are analogous to those presented in the table that presents the main estimate. Standard errors calculated by school cluster. * p<0.1, ** p<0.05, *** p<0.01.

Tables A.10 Models of quartiles on the initial scores distribution

Panel with fixed effects. Models of quartiles on the initial scores distribution, effect of SEN in Mathematics

	Q1	Q2	Q3	Q4
Class with TSEN integration (β_1)	-2.675***	-2.410***	-2.437***	-1.394***
	(0.699)	(0.566)	(0.520)	(0.556)
Class with PSEN integration (β_2)	-3.209***	-3.412***	-1.237*	-0.489
	(0.846)	(0.700)	(0.636)	(0.673)
TSEN*Time (β_3)	4.121***	1.415**	0.587	-0.699
	(0.754)	(0.622)	(0.570)	(0.616)
PSEN*Time (β_4)	2.840***	4.219***	1.545**	-0.690
	(0.936)	(0.767)	(0.705)	(0.755)
N	49,328	54,966	61,854	58,348

This table presents FE estimations for the 2007-2011 panel. Models Q1 (1-25%), Q2 (26-50%), Q3(51-75%) and Q4(76-100%), refers to samples of quartiles of 2007 scores distribution in math.
Household characteristics include controls that vary over time, such as books, income, and parents' schooling. School characteristics include controls by type (municipal, voucher, or private). Class characteristics comprise controls by parents' average schooling, class size, and evidence of SEN sorting. Standard errors calculated per individual cluster over time. * p<0.1, ** p<0.05, *** p<0.01.

Panel with fixed effects. Models of quartiles on the initial scores distribution, effect of SEN in Reading

	Q1	Q2	Q3	Q4
Class with TSEN integration (β_1)	-2.204***	0.274	-0.220	-0.782
	(0.684)	(0.602)	(0.557)	(0.587)
Class with PSEN integration (β_2)	-1.023	-2.608***	-1.803***	-0.526
	(0.841)	(0.777)	(0.691)	(0.733)
TSEN*Time (β_3)	1.968***	0.568	-0.07	0.592
	(0.744)	(0.664)	(0.617)	(0.659)
PSEN*Time (β_4)	0.286***	2.249***	3.994***	1.973**
	(0.931)	(0.841)	(0.757)	(0.814)
N	50,174	54,654	61,494	58,198

This table presents FE estimations for the 2007-2011 panel. Models Q1 (1-25%), Q2 (26-50%), Q3(51-75%) and Q4(76-100%), refers to samples of quartiles of 2007 scores distribution in reading.
Household characteristics include controls that vary over time, such as books, income, and parents' schooling. School characteristics include controls by type (municipal, voucher, or private). Class characteristics comprise controls by parents' average schooling, class size, and evidence of SEN sorting. Standard errors calculated per individual cluster over time. * p<0.1, ** p<0.05, *** p<0.01.

Tables A.11: FE models with restricted sample in school integration variation over time

FE model with restricted sample in school TSEN variation over time: effect of class with TSEN

	Mathematics			Reading		
	(1)	(2)	(3)	(1)	(2)	(3)
Class with TSEN integration	-2.024***	-1.770***	-1.711***	-0.683**	-0.0752	-0.0201
	(0.312)	(0.335)	(0.386)	(0.320)	(0.343)	(0.399)
TSEN*Time	1.462***	1.242***	1.319***	1.005***	0.794**	0.986**
	(0.331)	(0.351)	(0.391)	(0.341)	(0.362)	(0.404)
N	206,358	191,970	169,774	206,358	191,970	169,774

This table presents estimations for the 2007-2011 panel. Left panel models use mathematics SIMCE scores as the dependent variable. Right Panel models use reading SIMCE scores as the dependent variable. Models (1), (2) and (3), restrict the sample to <10% variation, <5% variation and <3% variation, respectively.
Household characteristics include controls that vary over time, such as books, income, and parents' schooling. School characteristics include controls by type (municipal or voucher). Class characteristics comprise controls by parents' average schooling, class size, and evidence of SEN sorting. Standard errors calculated per individual cluster over time. * p<0.1, ** p<0.05, *** p<0.01.

FE model with restricted sample in school PSEN variation over time: effect of class with PSEN

	Mathematics			Reading		
	(1)	(2)	(3)	(1)	(2)	(3)
Class with PSEN integration	-1.961***	-1.753***	-1.418***	-1.413***	-1.145***	-1.458***
	(0.381)	(0.405)	(0.454)	(0.395)	(0.420)	(0.476)
PSEN*Time	1.994***	1.486***	0.971**	2.089***	1.598***	1.755***
	(0.407)	(0.426)	(0.460)	(0.420)	(0.441)	(0.479)
N	208,716	199,970	186,074	206,358	191,970	169,774

This table presents estimations for the 2007-2011 panel. Left panel models use mathematics SIMCE scores as the dependent variable. Right Panel models use reading SIMCE scores as the dependent variable. Models (1), (2) and (3), restrict the sample to <10% variation, <5% variation and <3% variation, respectively.
Household characteristics include controls that vary over time, such as books, income, and parents' schooling. School characteristics include controls by type (municipal or voucher). Class characteristics comprise controls by

parents' average schooling, class size, and evidence of SEN sorting. Standard errors calculated per individual cluster over time. * p<0.1, ** p<0.05, *** p<0.01.

FE model with restricted sample in school SEN variation over time: effect of class with SEN

	Mathematics			Reading		
	(1)	(2)	(3)	(1)	(2)	(3)
Class with SEN integration	-2.257***	-1.917***	-1.747***	-1.106***	-0.509	-0.486
	(0.294)	(0.319)	(0.369)	(0.302)	(0.328)	(0.383)
SEN*Time	1.472***	1.396***	0.898**	1.076***	0.831*	0.661*
	(0.308)	(0.328)	(0.365)	(0.318)	(0.338)	(0.377)
N	202,698	182,242	155,294	202,698	182,242	155,294

This table presents estimations for the 2007-2011 panel. Left panel models use mathematics SIMCE scores as the dependent variable. Right Panel models use reading SIMCE scores as the dependent variable. Models (1), (2) and (3), restrict the sample to <10% variation, <5% variation and <3% variation, respectively.
Household characteristics include controls that vary over time, such as books, income, and parents' schooling. School characteristics include controls by type (municipal or voucher). Class characteristics comprise controls by parents' average schooling, class size, and evidence of SEN sorting. Standard errors calculated per individual cluster over time. * p<0.1, ** p<0.05, *** p<0.01.

Tables A.12: Panel with fixed effects – Sample without evidence of PSEN/TSEN sorting

Panel with fixed effects – Sample without evidence of PSEN/TSEN sorting

	Mathematics		Reading	
TSEN*Time (β_3)	0.823**	1.002***	0.506	0.525
	(0.328)	(0.349)	(0.337)	(0.359)
PSEN*Time (β_4)	1.806***	1.834***	2.272***	2.309***
	(0.400)	(0.425)	(0.413)	(0.438)
Class with TSEN integration (β_1)	-2.363***	-1.575***	-1.177***	-0.230
	(0.277)	(0.312)	(0.285)	(0.320)
Class with PSEN integration (β_2)	-1.476***	-1.528***	-1.454***	-1.173***
	(0.352)	(0.387)	(0.363)	(0.401)
Individual and time fixed effects (FE)	Yes	Yes	Yes	Yes
School FE	No	Yes	No	Yes
Family characteristics	Yes	Yes	Yes	Yes
Class and school characteristics	Yes	Yes	Yes	Yes
N	225,655	218,398	225,655	218,398

This table presents estimations for the 2007-2011 panel, but only considering the sample without evidence of sorting. Columns 1 and 2 use mathematics SIMCE scores as the dependent variable. Columns 3 and 4 use reading SIMCE scores as the dependent variable. Household characteristics include controls that vary over time, such as books, income, and parents' schooling. School characteristics include controls by type (municipal, voucher, or private). Class characteristics comprise controls by parents' average schooling, class size. Standard errors calculated per individual cluster over time. * p<0.1, ** p<0.05, *** p<0.01.