THE DANISH NATIONAL CENTRE FOR SOCIAL RESEARCH

Beatrice Schindler Rangvid

05:2016 WORKING PAPER

RETURNING SPECIAL EDUCATION STUDENTS TO REGULAR CLASSROOMS: EXTERNALITIES ON PEERS' OUTCOMES

SFI – THE DANISH NATIONAL CENTRE FOR SOCIAL RESEARCH

INKLUSIONSPANELET

RETURNING SPECIAL EDUCATION STUDENTS TO REGULAR CLASSROOMS: EXTERNALITIES ON PEERS' OUTCOMES

Beatrice Schindler Rangvid

THE DANISH NATIONAL CENTRE FOR SOCIAL RESEARCH, COPENHAGEN, DENMARK

Working Paper 05:2016

RETURNING SPECIAL EDUCATION STUDENTS TO REGULAR CLASSROOMS: EXTERNALITIES ON PEERS' OUTCOMES*

Beatrice Schindler RANGVID
The Danish National Centre for Social Research (SFI)

October 2016

Abstract

Policy reforms to boost full inclusion and conventional return flows send students with special educational needs (SEN) from segregated settings to regular classrooms. Using full population micro data from Denmark, I investigate whether becoming exposed to returning SEN students affects the academic achievement of other students in the school-grade cohort. The basic identification strategy controls for student and school-by-grade fixed effects in value-added test scores and is similar to the model in Hanushek, Kain & Rivkin (2002). I add a third dimension to disentangle differential exposure effects in reform years (=years with extraordinary large numbers of returners). Together with the basic model, this yields a triple differences model providing strong causal identification. The main finding is that becoming exposed to a returning SEN student during the reform period has a negative impact on test score gains of moderate size (-0.036 SD), while no significant effect is found in non-reform years. The results are robust to sensitivity checks. The negative exposure effect is slightly larger for boys, but does not differ by parental education or grade-level.

Keywords: education economics; fixed effects; difference-in-differences; triple differences; value-added; education; special needs education; mainstreaming; externalities; peer effects
JEL Classification: I20

* Contact information: bsr@sfi.dk; The Danish National Centre for Social Research (SFI), Herluf Trolles Gade 11, 1052 Copenhagen K, Denmark. This analysis is part of a project commissioned by the Danish Ministry for Children, Education and Gender Equality. The usual disclaimer applies. Thanks to Eskil Heinesen, Paul Bingley, Rodrigo Torres and participants at the EALE 2016 conference and workshops at IEB (Barcelona) and CEN (Copenhagen) for valuable comments and suggestions.

1. Introduction

Countries around the world organize education for children with special education needs (SEN) in different ways. The models adopted include segregated settings like special schools or separate classes in mainstream schools, and fully inclusive settings where SEN students are taught alongside their regular programme peers. Average per student costs in segregated settings tends to be significantly higher than in regular classes, but there is no conclusive evidence of the benefits of segregated education. Yet, in many countries, education opportunities for SEN students include teaching in segregated settings. As placement in special schools or classes may not necessarily last for the whole school career, some students eventually return to regular classrooms.

Another factor that may send students back to regular classrooms is a heightened political focus on inclusive education. Today, the general thrust is to provide special needs education in a fully inclusive setting (i.e. in regular classrooms), which is widely regarded as desirable for equality and human rights (World Health Organisation 2011). Yet, researchers, policymakers, practitioners, and parents have raised concerns about the impact that returning SEN students may have on the learning outcomes of their regular classroom peers. As returners are not used to coping in a regular classroom setting, the transition into regular classrooms is likely to pose challenges both to the returning student, but also to teachers and peers and may influence learning outcomes. Such externalities may come in the form of disruptions, changing pedagogy or altering resources in regular classrooms. Understanding these effects is therefore crucial for educational policy makers organising the transition towards more inclusive education.

Across European countries 2.3% of students within compulsory schooling are educated in segregated settings (World Health Organisation, 2011). Yet, countries vary widely in the numbers of children who receive education in segregated settings. In some countries segregated education is virtually non-existent (Italy, Portugal, Spain), while other countries exclude part of the student population from fully inclusive education (e.g. Belgium (French), Denmark, Finland, France, the Netherlands)[1]. Also in the US, raising the share of SEN students educated in regular classrooms is a policy target (Fletcher, 2010).

[1] The Netherlands, France and Finland educate 2-4% and Belgium (French) and Denmark 4-6% of their students in segregated settings (author's own calculations based on data provided in European Agency for Development in Special Needs Education, 2012). The data refers to the percentage of all compulsory-age students, not only to students who have been officially identified as having SEN. There are no universally agreed definitions for the concepts of special needs education. International comparison of data on children with special education needs is hampered by differences in definitions, classifications, and categorizations. The only comparable data is the percentage of students who are educated in segregated settings. The European Agency for Development in Special Needs Education has an operational definition for segregation (*education where the pupil with special needs follows education in separate special classes or special schools for the largest part (80% or more) of the school day*), which most countries agree upon and use in data collection.

While there is a general consensus that all learners should be educated in fully inclusive settings to the extent possible, the evidence on the impact of mainstream education on academic outcomes for SEN students is not conclusive, and there is very little empirical evidence on potential spillovers on classmates. Recent evidence on the effects of the inclusion of children with various kinds of disadvantages in regular classrooms suggests negative spillover effects on their peers[2]. However, research focusing specifically on the effect of including SEN students in mainstream classrooms on their peers' outcomes is scarce and results are mixed. Estimating the effect of having SEN students as class mates, Fletcher (2009, 2010) and Gottfried & Harven (2015) find negative effects on other learners.[3] Yet, when taking account of within-grade sorting across classes by estimating same-grade peer effects Hanushek, Kain and Rivkin (2002) and Friesen et al. (2010) find no significant effects.

This study uses full population data from Denmark to examine effects on peers associated with returning SEN students from segregated education to regular classrooms. Specifically, I estimate the impact on achievement gains of becoming exposed to a same-grade SEN peer who has recently returned from segregated education to a regular classroom. The model differentiates between exposure effects during non-reform years (with normal return flows) and effects during reform years with large return flows caused by policy efforts to increase the percentage of students educated in fully inclusive settings. Using a value-added approach, achievement is effectively measured by test score gains. Moreover, to control for nonrandom selection, this study takes advantage of longitudinal information from administrative registers on individual students in multiple student-grade cohorts to control implicitly for time-invariant student and school-by-grade effects on test score gains. This is similar to the specification in Hanushek, Kain, Rivkin (2002), but is here applied to exposure to recent returners (instead of to SEN students in general). The effect of becoming exposed to a returning SEN student on the achievement of regular classroom peers is thus identified by changes over time in the presence of a returner in one's grade-cohort. To disentangle the differential exposure effects in reform years (i.e. years with extraordinary large numbers of returners), I enhance the main specification by another dimension. This *de facto* triple differences model provides strong causal identification.

Denmark is an excellent place to study effects of returning students. First, a substantial degree of openness between different educational settings gives rise to flows between inclusive and segregated educational settings. On top of this, the policy effort to increase inclusion rates boosts the number of returning SEN students in the reform period. Starting from a situation with comparatively low inclusion rates in 2011 (94.4%), state policies have been directing schools to mainstream SEN students into

[2] See Carrel & Hoekstra (2010) on children exposed to domestic violence; Figlio (2007) and Kristoffersen et al. (2015) on disruptive children; Cho (2012) and Diette & Oyelere (2014) on non-native English speakers; Gottfried (2013) on grade-retained peers; and Gottfried (2014) on tardy class mates.
[3] Yet, the effect on reading in Fletcher (2010) is only marginally significant at the 10% level.

regular classrooms. A national target has been set to increase the inclusion rate to 96% within a three year period (2012-2015). Reaching this goal both implies keeping more students in inclusive settings in the first place, but also bringing back students from segregated settings to regular classes.[4] Policymakers and parents in the schools that take in these students may wonder what the consequences will be for the other learners in the receiving classrooms. This study focusses on the impact of the latter on peer outcomes in receiving cohorts.

This study contributes to the literature as the first study of the externalities of returning SEN students on their peers' educational outcomes. Returners may be more difficult to accommodate in regular classes and may be more disruptive[5] than SEN students who were never educated in segregated settings – both because they may on average have a higher degree of special needs (as they have been sent to special classrooms and schools in the first place), but also because they face an entirely new education setting upon their return that they may not easily adapt to. Understanding the impact of returners is relevant, in particular for school systems moving towards a more inclusive approach to educating students with special needs. On top of contributing to the literature on the effects of SEN students in regular classes, this study places itself within the broader peer effects literature (Epple and Romano (2011) and Sacerdote (2011) provide overviews). For example, linking peer achievement or characteristics to students' learning, contributions related this study are Ammermueller & Pischke (2009), Angrist & Lang (2004), Hoxby & Weingarth (2007), Lavy, Paserman & Schlosser (2012), Lavy, Silva & Weinhardt (2012), Lefgren (2004) and McEwan (2003).

The main results provide evidence of negative externalities of returning SEN students on peers' reading gains during reform years with large return flows of SEN students. The negative effect is of modest size though (-0.036 SD). In non-reform years, there is no evidence of negative externalities on peers. The exposure effect during reform years is negative and significant only for boys, while I do not find differential effects by parental education or grade-level.

The structure of the paper is as follows: Section 2 provides a short presentation of the background and institutional setting and gives a brief description of the group of returning students. Section 3 presents the identification strategy, and Section 4 introduces the data. Section 5 reports the results and Section 6 offers concluding remarks.

[4] This means that students attend a regular classroom (in a mainstream school) for the main part of the school week. Students may receive extra support if needed either within the regular classroom or in separate resource rooms.

[5] Lazear (2001) presents a theoretical framework linking class room disruptions to learning.

2. Background

Historically, children with special educational needs have often been excluded from mainstream education opportunities. In many countries early provision of education was generally through separate special schools. In 1994 the World Conference on Special Needs Education in Salamanca, Spain, produced the so-called Salamanca Declaration, a statement and framework for action, which encouraged governments to design education systems that provide for groups with diverse needs, so that all students can have access to regular schools (World Health Organization, 2011). Moreover, over the last decade, a general consensus has developed to move from a broad sense of inclusion, whereby education may take place in a range of settings, such as special schools, or special classes, or regular classes in mainstream schools, to a stricter sense of inclusion, under which all children with disabilities should be educated in regular classrooms with age-appropriate peers. Today, many school systems hold both inclusive and segregated alternatives for SEN students. By international comparison, Denmark has a rather large share of students educated in segregated settings. Moves between segregated settings and regular classrooms are common due to frequent reevaluations of SEN students' development, which makes Denmark a natural choice when studying externalities of returning students to regular classrooms.

Half a decade ago, rapidly rising spending on special needs education in Denmark and high exclusion rates highlighted the need for reform with the aim to provide education for more students in regular classrooms. Average per pupil costs in segregated settings is significantly higher than for students in regular classes.[6] In 2008/09, segregated special needs education absorbed more than 80% of overall spending on special needs education (Ministry of Finance, 2010). In 2012, a reform of special education was implemented including the target to increase inclusion rates from 94.4% to 96% by 2015. A legislative change in the Folkeskole Act (narrowing the definition of special needs education to include only extensive support) and an agreement between the municipalities and the national government (outlining objectives for increased inclusion) were followed by a change in the economic incentives for the schools to include students with special needs education in regular classes. This was accomplished by decentralizing financial responsibility for special needs education from the municipal level to the schools. This change gave strong financial incentives for the schools to provide for SEN students in inclusive settings (as segregated settings are considerably

[6] According to Ministry of Finance (2010, p. 15) in 2008/09 costs averaged DKK 280,000 – or $45,000 - for students in special schools, DKK 185,000 – or $30,000 - for students in special classes, compared to DKK 85,000– or $14,000 - for SEN students in regular classes. These numbers are not directly comparable, since the educational needs of SEN students who are included in regular classes are different that those of students educated in segregated settings. Yet, this example highlights a large variance in costs. To the extent that these may not (all) be justified for (all) students, money might be more effectively spent otherwise.

more expensive).[7] Higher inclusion rates may be achieved both by (i) keeping more students with special needs in regular classes instead of segregating them to special schools or classes and by (ii) returning students from segregated settings to regular classrooms. Evidently, both measures imply that more regular students than before have SEN-students as classmates, including recently returned students. Of both measures, returning students from segregated settings to regular classrooms was by far the most controversial part of the reform and has caused concern among parents and teachers in regular classrooms.

In order to meet the challenge of moving towards more inclusive education, a number of initiatives have been undertaken to support the reorganization process in the municipalities including financial incentives, school reform[8], information and attitude campaigns for parents and students and follow-up work.[9] Initiatives to make regular classrooms more inclusive include a strengthened focus on individualized teaching in regular classrooms, temporary subdivision of class, additional lessons, two teachers in class, teachers' assistants, and individual support to help the SEN students overcome practical obstacles related to school attendance.

Special education in Denmark

In Denmark, special education may take place in a range of settings – such as special schools, special classes or regular classes in mainstream schools. In 2013, Denmark adopted a narrow definition of special needs education, which includes only students with more than 9 hours a week (or 12 teaching hours) of extra support. In 2015, five percent of all students in public schools are defined as having SEN.[10] Only 5% of these students are mainstreamed in regular classes, while the remaining 95% are taught in segregated settings (39% in special schools and 56% in separate classes in mainstream schools). Overall, in 2015, 4.8% of all students are taught in segregated settings. Thus, the inclusion rate – the percentage of students educated in mainstream classrooms - is 95.2% in 2015 (up from 94.2% in 2011).

Since the change in the definition, the number of SEN students with less than 9 hours of special needs education is unknown, because data on special education services relating to these students is no longer collected. The most recent number available is from 2012.

[7] Research has shown that changes in fiscal incentives are related to changes in special education growth in the US (Cullen 2003, Kwak 2010). Furthermore, a rise in special education enrolment has also been related to the introduction of school accountability policies (Jacob 2005). In Denmark, the publication of school results was introduced during the period of rising exclusion rates. However, whether or not there is a causal relation remains to be examined.

[8] E.g. improving teachers' and staff's skills/professional development, counseling and renewal of guidelines.

[9] E.g. monitoring of shift towards more inclusion by the Ministry, and other governmental support.

[10] Author's own calculation based on administrative data from the Ministry of Education (2015).

In 2012, 4.8%[11] of all compulsory-aged students received less than 9 hours extra support due to their (minor) special educational needs. Students with minor special needs remain in regular classrooms and receive extra support as a supplement to general teaching.

The proportion of compulsory-aged students who are identified as having SEN varies across countries, but international meaningful comparisons are hampered by differences in definitions and assessment of SEN-students. According to the (tentative) numbers provided by the European Agency for Development in Special Needs Education (2012), the percentage classified as SEN ranges from 1.6% of all compulsory aged students in Sweden to 24.2% in Iceland. Numbers for countries outside Europe are not provided in the report, but other sources show that the percentage in the US is about 13% (age 13, 2004) and 9% in British Columbia/Canada (grade 7, 2002-04)[12].

Returners

From 2012 to 2015, the percentage of students educated in segregated settings has decreased from 5.8% to 4.8%. Thus, more students are now educated in inclusive settings, either because they have avoided being segregated in the first place, or because they have returned from special schools or classes to regular classrooms. This study focuses on returners to regular classrooms.

Table 1: New returners as a percentage of public school students by grade level and schoolyear

	All grade levels	*Grade 0* *	*Grade 1*	*Grade 2*	*Grade 3*	*Grade 4*	*Grade 5*	*Grade 6*	*Grade 7*	*Grade 8*	*Grade 9*	*Grade 10* **
2009/10	0.29	0.03	0.22	0.21	0.24	0.18	0.36	0.24	0.31	0.30	0.32	1.62
2010/11	0.34	0.05	0.34	0.25	0.34	0.27	0.39	0.30	0.31	0.31	0.38	1.87
2011/12	0.54	0.06	0.40	0.68	0.51	0.51	0.53	0.47	0.54	0.51	0.74	2.44
2012/13	0.49	0.06	0.25	0.44	0.32	0.50	0.49	0.60	0.53	0.55	0.62	2.67
2013/14	0.33	0.03	0.15	0.19	0.27	0.25	0.27	0.28	0.34	0.36	0.44	2.82
2014/15	0.23	0.02	0.06	0.10	0.12	0.14	0.13	0.17	0.21	0.29	0.40	2.80

** These are the (few) students who attended grade 0 in a special school or special class and then proceeded to regular class, yet repeated grade 0 (in regular class).*
*** Grade 10 is a voluntary grade level, among public school students often used as an additional year of preparation before proceeding to upper secondary education.*

Source: Author's calculations using administrative microdata.

In 2009/10, 0.29% of all public school students were students who had returned from a special class or school to regular classrooms in the current school year (Table 1). This share increased to 0.54% in 2011/12, and fell to 0.23% in 2014/15, which by now is the last year of data available. The decreasing number of returners in recent years probably

[11] Author's own calculation based on administrative data from Statistics Denmark and Ministry of Education (2015).
[12] Friesen et al. (2010)

reflects the shrinking pool of students in segregated settings capable to return to regular classrooms. While there is always a natural flow into and out of segregated special education, the temporary rise and fall in return rates during our period of study reflects the policy reform towards more inclusive education. Overall, the share of returners is highest in grade-level 10 and lowest during the first years of primary education. Additional calculations show that returners are much more likely to come from special classes in mainstream schools than from special schools. Only 16% of returners attended a special school the year before return, while 72% of returners attended a special class in a mainstream public school.[13] 82% of the returners in the relevant grade-levels have taken the national test in reading the year before they return – a substantially higher share than the 30-50% among all students in segregated education[14] and much closer to the share of 92% among regular program students. Thus, the share taking the reading test is considerably higher among returners indicating that returners - on average – have milder impairments than stayers. [15]

Upon return, two out of three returners attend regular classes in public schools – either regular public schools (61%) or youth schools (*ungdomsskoler*) which only offer grades 8-10 (7%). 24% return to private continuation schools (*efterskoler*[16]) and 7% return to other private schools. The large majority (86%) are placed in the subsequent grade-level, while others repeat a year. Only 13% receive extensive extra support (i.e. more than 9 hours a week) when returning to a regular classroom. Overall, regular classroom peers in receiving schools are somewhat more disadvantaged than students in non-receiving schools, but the differences are small: reading scores for receiving peers are 0.025 SD lower than for non-receivers, the share of immigrant students is 8.3% in receiving schools and 6.7% in non-receiving schools, and the share of mothers with no more than compulsory education is 20% in receiving schools compared to 18% in non-receiving schools.

3. Estimation Strategy

The empirical analysis examines the effect on achievement gains for regular-education students of being exposed to a *recent returner* in their school-cohort. In this study, I

[13] Before return, 6% attended school at a day-treatment facility or a live-in treatment facility. The remaining returners come from special continuation schools (2%; efterskoler med særligt tilbud), special classes in youth schools (2%), special classes in private schools (1.5%) and the remaining 0.5% return from other school types.

[14] The share taking the national test varies by grade-levels and type of setting (special school or classes). Participation is generally higher among older students and among students in special classes (Rangvid & Lynggaard, 2014).

[15] Moreover, among those who take the tests, returners score higher than stayers.

[16] As an alternative to the lower-secondary education (grades 8, 9 and 10) at the *folkeskole*, students have the opportunity to attend residential so-called continuation schools (efterskoler) from the ages of 14 – 18 for one or two years.

refer to students who returned from segregated education to a regular classroom during the current or the previous school-year as *recent returners*.

A challenge for identifying causal effects arises if the probability of becoming exposed to a recent returner is related to unobserved student and school characteristics, for example if principals match returning student to teachers and peers based on unmeasured characteristics that also affect achievement. This paper exploits the panel nature of the dataset with multiple gain scores for each student and multiple grade cohorts for each school to control for much of the confounding variation. The effect of externalities is estimated in a value-added framework with student and school-by-grade fixed effects, allowing for systematic, but unmeasured differences in students and schools.[17]

As a baseline for the analysis, a value-added model is estimated. Value-added models are generally superior to simple cross-sections, because using the gain in achievement between two test years as the outcome removes any fixed student effects on the level of achievement. As a result, the condition for causal identification is less restrictive, since exposure to recent returners needs only be exogenous to *changes* (not *levels*) in test scores.

The equation

$$Read_{igst} = \alpha + \beta Read_{igs,t-1} + \delta X_{igs,0} + \theta RR_{gst} + \nu_t + \varphi_{gs} + \varepsilon_{igst} \qquad (1)$$

models the reading test score, *Read*, for student *i* in grade *g* and school *s* at time *t* as a function of vectors of student and family characteristics (*X*), exposure to a recent returner in the grade-cohort (*RR*), and two error components: a year fixed effect ν_t, and a random error (ε_{igst}). The student's reading score in period *t-1*, $Read_{igs,t-1}$, captures the accumulated knowledge or skills students have at the beginning of the between-test period. Family characteristics, $X_{igs,0}$, are included because they may be correlated with achievement growth. They are measured at age five, the year before the official age for starting school. The coefficient θ captures the impact of having a recent returner in the same grade-level on reading test score gains of regular students. However, even when controlling for observable student and family differences in achievement gains, any differences across schools and across grades within schools that are not perfectly correlated with the student fixed effect in gain scores, but correlated with the probability of becoming exposed to a recent returner in one's grade cohort would still bias the results. Including school-by-grade fixed effects, φ_{gs}, exploits variation over time in

[17] Preferably, one would include school-by-grade-by-year fixed effects to also allow for time-varying differences across grades within schools. However, this is not feasible since the primary variable of interest – the presence of a recent returner in the school-grade cohort - is measured at that level. Instead, a robustness check using school-year fixed effects is provided as are results from an estimation including characteristics of the school-grade-year cohort.

exposure to a recent returner within schools and grade-levels, and controls for much of the confounding variation across schools and grades.[18]

However, equation (1) takes only account of observable differences in student background. If there are unobserved differences in students learning ability, multiple gain score records per student would control for such differences. Therefore, to account for all time-invariant individual differences in *learning* across students, the model is reestimated with two records of gain scores for each student and including student fixed effects. However, to avoid the problems inherent in dynamic panel data models (Angrist & Lavy, 2009), I use a common alternative specification of the value-added model in difference form, where the dependent variable is the difference between test scores and lagged test scores. The model in difference form is more restrictive since it effectively restricts the coefficient of the lagged score, β, to equal 1 (Todd & Wolpin, 2003).[19] To evaluate the potential effects of this restriction on the results, I repeat the analysis fixing β to alternative values. The results are similar.[20]

Formally, the equation can be written as

$$\Delta\, Read_{igst} \equiv Read_{igst} - Read_{igs,t-1} = \alpha + \theta RR_{gst} + \nu_t + \varphi_{gs} + \sigma_i + \varepsilon_{igst} \qquad (2)$$

where σ_i are student fixed effects.[21] This is the main model specification for the empirical analysis. In the analyses, I consider only *entries* into exposure, since exits may be less exogenous. Thus, $RR_{gst} = 0$ in the first gain-score spell for all students in the estimation sample, while RR_{gst} may equal 0 or 1 in the second period.[22]

Eliminating student fixed effects in gain scores means that θ in equation (2) is identified by the change in gain scores for students who are not exposed to a recent returner during the first period, but are exposed in the second period. The estimate for θ now captures the effect of idiosyncratic year-to-year variation in becoming exposed to a recent returner within the school and grade level. Thus, in the fully specified model, the effects of externalities induced by recent returners are identified as the difference in gain scores for students who transit from non-exposure to exposure to a recent returner on the one hand and those who are not exposed in either period on the other hand - allowing for differences across grades and schools in teacher and school quality. Of the 123,378 students in the estimation dataset who are not exposed during their first two-year spell, 32,455 students – or 26% – are exposed in their second spell. The fact that a significant

[18] On top of this, any variation in school quality in grades over time that is related to the probability of becoming exposed to a recent returner would continue to bias the estimated coefficients. Yet, school-by-grade-by-year fixed effects cannot be included, because all students attending grade *g* in school *s* in year *t* experience the same exposure to recent returners. As part of the robustness analysis, I present results that include school-by-grade-by-year controls. The results are virtually identical (Table 5).

[19] This is equal to the model specification used in Hanushek, Kain & Rivkin (2002).

[20] The results of this robustness check are presented in Table 3.

[21] Note that time-invariant characteristics at the student level, $X_{igs,0}$, cancel out due to the inclusion of student fixed effects.

[22] An addition analysis shows no significant effects of *exiting* exposure.

number of peers switch from being not-exposed to being exposed strengthens the identification strategy.

Finally, I enhance the model to investigate the effect of exposure specifically during reform years. The variation that drives identification in the value-added student fixed effects specification derives both from the natural flow of returners to regular classrooms and from reform driven variation due to the policy effort to boost inclusion, which has led to an extraordinary increase in the number of returning students.[23] During reform years, an additional number of students must be included in regular classrooms which may put an extra strain on receiving schools. The cohorts in our study that are followed from 2010 to 2014, experience a larger number of returning students during their second two-year period between tests (2012-2014) than the cohorts observed between 2011 and 2015. While the total number of returners experienced in the first case is roughly 9,000, the equivalent number in for the other cohorts is only 6,700 (or 25% fewer returners).

To single out the specific effect of becoming exposed during years with considerably higher numbers of returners, I enhance the basic model specification with another dimension by adding an interaction term between exposure and an indicator that is equal to one for cohorts with potential treatment spells (i.e. the second gain score spell) during reform years. The estimate for the main effect provides the estimated exposure effect in reform years, while the interaction term captures the difference between the exposure effects in reform years and other years. The resulting specification is de facto a triple difference model, since comparing the difference between reform and non-reform years adds a third difference to the model. The enhanced model thus compares student outcomes across three differences (exposed/not exposed, gain score spells 1 and 2, reform/non-reform years) providing particularly strong causal identification.

The fully specified main specification thus becomes:

$$\Delta\, Read_{igst} = \alpha + \theta RR_{gst} + \eta RR_{gst} \times NR_c + \gamma NR_c + \nu_t + \varphi_{gs} + \sigma_i + \varepsilon_{igst} \qquad (3)$$

where NR_c indicates that the gain score spell falls in the non-reform period and $RR_{gst} \times NR_c$ is the interaction between exposure and non-reform years. The coefficient η is thus the difference in the exposure effect in non-reform years (compared to reform years). The estimates of main interest for our analysis are the exposure effect in reform years, θ, and the difference between reform and non-reform years, η.

Following the baseline analysis of average effects of exposure to recent returners for all students, I proceed by investigating the possibility that groups of students are affected differently. For example, boys may be affected in other ways than girls through social interactions with recent returners, because most recent returners are boys (70%), or externalities may differentially affect students from different family backgrounds.

[23] The number of returning students will only be higher for a shorter period, since the policy also implies fewer student being segregated in the first place – and therefore, the flow back from segregated education will decrease as well after a transition period.

A potential threat to the identification strategy would arise if mainstream peers selectively opt out of exposed grade-levels. However, as long as these students take the exams at their new school, the intention to treat effect can still be measured. In the present case, this holds for students who switch to other public schools, since all public school students are assessed by the national tests. Post-outcomes for these students will therefore be measured even if students opt out and they will be included in the analysis thus avoiding sample selection. However, students who opt out of public school and enroll in private schools will no longer be assessed by national tests, and this could bias the results.[24] Yet, auxiliary analyses show that the share of peers with missing post-scores is similar among exposed and unexposed students and thus, this kind of sample selection does not seem to be a major concern in this study.[25]

4. Data

This study utilizes administrative microdata to shed light on the peer effects of returning SEN students from segregated settings to regular classrooms[26]. These data track successive cohorts of students as they progress through school and contain extensive and reliable information on test scores and students' family background, as well as school and grade-level identifiers. Full population data is crucial for this analysis, because only a small number of students move from segregated settings to regular classrooms each year.

Unique IDs permit linking the student records with separate special-education information on academic setting for SEN students, ranging from assistance while in the regular classroom over separate classes in regular schools to separate schools. Importantly for the empirical strategy, the data allows us both to identify SEN students who move from segregated to inclusive educational settings, and to select same-cohort peers of these returners.

Although the administrative data holds much information on education careers for many student cohorts back in time, data for test scores which are essential for this analysis are available only from the schoolyear 2009/2010. Information on the educational setting of special needs students is available from 2008/09.[27] As a result, key data needed for the present study is currently available for the schoolyears 2009/10 to 2014/15.

At the compulsory school level, there are roughly 65,000 students in each grade-cohort. 55,000 of these attend one of 1,250 regular public schools[28], totaling 550,000 students

[24] Administering the national tests is voluntary for private schools. Only few do.
[25] The shares of students with missing post-scores among those exposed in grades 4, 6 and 8 is 0.0005, 0.018 and 0.068, respectively. The corresponding shares for the non-exposed are 0.0007, 0.020 and 0.076.
[26] Data is cleaned and stored at Statistics Denmark.
[27] In the data available to researchers, this information is available only from 2010/11. However, notably for this analysis, the Ministry of Children, Education and Gender Equality has made available data for earlier years.
[28] This definition includes the so-called *folkeskoler* and youth schools (*ungdomsskoler)*.

in ten compulsory grade-levels each year. For the estimations, I only keep students in regular classrooms which decreases the sample size to roughly 535,000 students per year.

Table 2: Structure of dataset

	Grade 0	Grade 1	Grade 2	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10
2009/10			X		X						
2010/11			X		X						
2011/12					X		X				
2012/13					X		X				
2013/14							X		X		
2014/15							X		X		

Note: The four different colours denote the four student cohorts used for the empirical analysis. The Xs mark grade levels and years with reading assessments.

Furthermore, for the main estimation sample student cohorts that are observed in three subsequent test-grades are selected (i.e. either grades 2, 4 and 6 or grades 4, 6 and 8[29]). Four cohorts meet this requirement (Table 2). I use data for grade-levels two through six for the two younger cohorts and grade-levels four through eight for the two older cohorts, about 203,000 students. The oldest cohort attended sixth grade in 2011/12 and the youngest cohort attended sixth grade in 2014/15. Table 2 shows the structure of the data.

The estimation sample is further restricted to students who have non-missing information on all three test scores[30], reducing the sample to about 173,000 students. Last, we only consider *entry* into exposure, because exits might be less exogenous. This reduces the final estimation sample to about 123,000 students.

The primary variable of interest is exposure to a recent returner, i.e. the presence of a recent returner in the same grade, school, and year in at least one of the two years between the pre-testscore and the post-testscore in the treatment period. Exposure is measured at the school-cohort level rather than at the class level due to potential sorting of recent returner across classrooms within a school and grade based on unobservable criteria. Recent returners are defined as students who have been educated in segregated settings (special school or special classes), but have returned to inclusive education (regular classes) in the current or the previous schoolyear. As in the overwhelming majority of cases, there is either none or only one recent returner in the school cohort,

[29] Reading tests are administered bi-annually each spring to students enrolled in grade-levels 2, 4, 6, and 8 in public schools.

[30] Furthermore, recent returners are excluded from the estimations as are a small number of school-grade-year cohorts due to suspected problems with data-quality for the information on educational setting (i.e. regular class, special class, and special school) in particular in the earlier years of data.

the variable of interest (exposure) is created as a binary variable indicating whether a student has at least one recent returner in the same grade-cohort or not.[31]

The main outcome variable is reading test scores from the national tests.[32] Beginning in 2010, reading tests were administered bi-annually each spring to students enrolled in grade-levels 2, 4, 6, and 8 in public schools, creating a two year gap between assessments. These mandatory tests are high-profile tests. They are IT-based and adaptive, meaning that tests are taken online at computers with the test system choosing questions based on the student's level of proficiency as displayed during the test and automatically calculating test results. The tests simultaneously evaluate the skill levels within three profile areas of reading: language comprehension, decoding, and reading comprehension. For the analyses in this paper, test scores have been standardized for each test, grade and year to have mean zero and standard deviation one using children in the entire sample.[33]

The subsequent empirical analyses focus on the effect on regular programme peers of being exposed to returning SEN students and thus rely heavily on transitions out of special education. In the estimation dataset, on average 26% of students switch from not being exposed in the first period to being exposed in the second period. Among the students with treatment spell in the reform period, 31% switch to being exposed compared to only 22% of those with treatment spell in the non-reform period.

5. Results

Table 3 presents estimated effects of becoming exposed to a recent returner in the same-grade cohort. Results from five different specifications are presented: the uncorrected correlation (1), a baseline value-added model (2), and the main specifications that also includes student fixed effects and school-by-grade fixed effects (3). In the main specification, the coefficient of the lagged reading score, β, from equation (1) is fixed to 1. Specifications 4 and 5 evaluate the effect of this restriction on the results by fixing β to alternative values (0.5; 0.75).

In addition to the indicator for the presence of a recent returner in the school-cohort (exposure term), all regressions include time fixed effects and models 1 and 2 also include grade fixed effects. Importantly, all regressions include an interaction term

[31] Additionally, I have considered including an indicator for whether a school cohort receives any new student during the two-year spell to control for potential disruption from any new students, not only from returners. However, a look at the data showed that this cannot be reliably estimated due to insufficient variation in the variable since 99% of all school cohorts receive new students over a two-year period.

[32] In addition to reading tests, we would like to consider the effect on math as well. However, with the advanced identification design used in this study (value-added student fixed effects), three test scores per student are needed – yet, math is assessed only twice (in grade levels 3 and 6).

[33] Specifically, first the scores are standardized for each profile area for each grade-year combination. Then, scores are averaged across the three profile areas before I standardize the average for each grade-year combination. The resulting final measure of the reading score has a standard deviation of one and mean zero.

between exposure and an indicator for whether the treatment period falls in the non-reform or reform period (and the main effect of the latter). Models 2-5 include students' lagged reading scores.[34] Measured student background characteristics that do no vary over time drop out in model 4 due to the inclusion of student fixed effects. Both model 4 and 5 contain school-by-grade fixed effects.

Table 3 shows the results of the variables of main interest: the exposure effect during the reform years[35] and the difference in the exposure effect to other years (i.e. the interaction term). The third row shows the sum of the main exposure effect and the interaction, which is the estimate of the exposure effect in non-reform years. A test of significance of this sum tests for the existence of an exposure effect in non-reform years.

Table 3: Main regression results. Effect of exposure to recent returners on peers' reading score gains.

	(1)	(2)	(3)	(4)	(5)
			Main specification	***Robustness, main specification***	
	Simple model	Value-added (VA), lagged control	VA(LHS: change) + StudFE + School/grade FE (β=1)	*VA(LHS: change) + StudFE + School/grade FE (β=0.75)*	*VA(LHS: change) + StudFE + School/grade FE (β=0.5)*
(1) Exposure effect (reform years)	-0.1064*** (0.0181)	-0.0528*** (0.0122)	-0.0365(*) (0.0198)	*-0.0358 * (0.0173)*	*-0.0352 * (0.0152)*
(2) *Interaction:* Additional exposure effect for non-reform years	0.0859** (0.0283)	0.0498** (0.0191)	0.0082 (0.0278)	*0.0117 (0.0244)*	*0.0151 (0.0217)*
(3) [=(1)+(2)] Exposure effect (non-reform years)	*-0.0205*	*-0.0030*	-0.0283	*-0.0241*	*-0.0201*
Year fixed effects	x	x	x	*x*	*x*
Lagged reading scores		x	x	*x*	*x*
Student fixed effects			x	*x*	*x*
School-by-grade fixed effects			x	*x*	*x*

*(*) p<0.10, * p<0.05, ** p<0.01, *** p<0.001 # observations: 246,756; # students: 123,378*
Standard errors are reported in parentheses and clustered by grade/school/year level.

The results in Table 3 show that the simple correlation of exposure and test scores during reform years is negative and highly significant (model 1). Including lagged testscores in the value-added specification (2) reduces the point estimate by half suggesting that exposed students have somewhat lower reading scores prior to exposure. However, while reduced, the value-added estimate is negative and highly significant. The main model specification (4) with student fixed effects and school-by-

[34] Note that the lagged reading score is entered as a control in model 2, while it is part of the test-score change dependent variable in models 3-5.
[35] The indicator is equal to 0 in the non-reform period and equal to 1 in the reform period. Thus, the exposure (main) effect is the effect during reform years and the interaction effect is the additional effect in non-reform years.

grade fixed effects further reduces the point estimate of the exposure effect. The estimate is significant at the 6% level and the effect size is -0.037 SD. The results of the robustness checks in specifications 4 and 5 confirm those of the main model specification: the estimates are similar (-0.035 and -0.036 SD) and are significant at the 5% level. Together, these results suggest that becoming exposed to recently returned students during reform years has a negative effect on achievement gains. However, the size of the effect, a decrease of roughly 0.037 SD of the reading score distribution, is moderate and approximately equivalent to half a month of learning gains per year.[36]

The positive sign of the interaction between exposure and the reform-years indicator suggest that the exposure effect is closer to zero in non-reform years. The third row shows the sum of the main effect and the interaction effect, which gives the exposure effect in non-reform years. While the point estimates are negative for all specifications, they are small and not significantly different from zero. Thus, the overall picture is no exposure effects during non-reform years.

Since this is the first study examining the impact of mainstreaming SEN students who previously were educated in segregated settings, the results cannot be directly compared to existing work. However, the zero results in non-reform years in this study corroborate results from related research on the effect of exposure to SEN students (i.e. not necessarily recent returners) suggesting that there are no adverse effects of a higher share of SEN students in the grade-cohort (Hanushek, Kain and Rivkin, 2002); Friesen et al., 2010). Also Fletcher (2010) finds no significant effects for reading scores, but negative effects for mathematics.[37] The negative effects during reform years found in this study are confirmed by another Danish study that mainly covers the reform period and examines the effect of exposure to children with psychiatric disorders (Kristoffersen et al., 2015). Kristoffersen et al. find negative effects on peers' reading scores of similar size (-0.02 SD) as in this study.

I run a series of robustness checks to examine the sensitivity of the results. Table 4 presents results (column 1 repeats the main results from model 4, Table 3, for comparison purposes). First, instead of including school-by-grade fixed effects, ideally one would include school-by-grade-*by-year* fixed effects to control for school-by-grade differences that vary over time. This is not possible, however, because the variable of interest – exposure to a recent returner – is measured at this level and could not be separately estimated if school-by-grade-by-year fixed effects were included as well. Instead, I add characteristics at the school-by-grade-by-year level to check the robustness of the main specification. The results in Table 4, column 2, show that

[36] This is an estimate for reading scores for 6th grade students (Andersen et al., 2014).

[37] The effect on math scores cannot be investigated with the strong identification strategy used in this study, because to produce two gain score spells, three test score measurements per student are required. However, unlike reading tests math tests are administered only twice (in grade levels 3 and 6).

including measures of the student body composition at the school/grade/year level[38] leaves the results virtually unchanged.

Table 4: Regression results. Robustness checks.

		(1)	(2)	(3)
		Main model results, repeated	Main model + school/grade/year controls	Main model with school-clustered standard errors
Main specification (model 4, Table 3)	(1) Exposure effect (reform years)	-0.0365(*)	-0.0367(*)	-0.0365
		(0.0198)	(0.0198)	(0.0276)
	(2) *Interaction:* Additional exposure effect for non-reform years	0.0082	0.0091	0.0082
		(0.0278)	(0.0278)	(0.0372)
	(1)+(2) Exposure effect (non-reform years) [2]	-0.0283	-0.0276	-0.0283
Alternative specification (model 5, Table 3)	(1) Exposure effect (reform years)	-0.0355**	-0.0354**	-0.0355*
		(0.0120)	(0.0120)	(0.0150)
	(2) *Interaction:* Additional exposure effect for non-reform years	0.0123	0.0125	0.0123
		(0.0170)	(0.0170)	(0.0220)
	(1)+(2) Exposure effect (non-reform years) [2]	-0.0232(*)	-0.0229(*)	-0.0232

Note: # observations: 246,756; # students: 123,378

*(*) p<0.10, * p<0.05, ** p<0.01, *** p<0.001 Standard errors are reported in parentheses, clustered at the grade/school/year level in models 1-3, and clustered at the school level in model 4.*

Second, the main model is re-estimated with standard errors clustered at the school level, instead at the school-by-grade-by-year level. Clustering at the school level is relevant if test scores are correlated within the school due to within-student correlation in test scores over time (and thus over grade levels). Using cluster correction at the more aggregate level reduces the levels of significance of the exposure estimates (reform years), but the results of the alternative specification are still significant at the 5% level. Overall, the results from the sensitivity analyses largely corroborate the results from Table 3.

Additional analyses

The analysis proceeds by exploring heterogeneous effects, because effects may vary for different groups of students. For example, exposure to a returner may affect boys more than girls through a same-gender peer effect, because most returners are boys. Or, the challenge of including a returner in a regular classroom may vary by student age, because return might be easier to cope with while students are younger than in later years. Furthermore, the effect of exposure may vary by the educational background of the peers' parents, because well-educated parents may be better able to support their

[38] Share of peers with unskilled mothers, share of native peers, and share of peers with mothers on social transfer payments.

children academically and push for additional support in the classroom if necessary. Thus, we expect the negative exposure effect to be smaller for children with well-educated parents.

To explore the possibility of heterogeneous effects for different student groups, the main specification has been estimated separately by gender, parental education[39] and grade level[40]. Table 5 provides results. The results suggest that the effect of being exposed to recent returners is marginally significant only for boys, suggesting that boys are harmed more by exposure to recent returners than girls. Furthermore, the exposure effect is marginally significant only for students with mothers without some tertiary education, but the point estimate is similar or even smaller to that for students with mothers with some tertiary education. The point estimate for older students is larger than for the younger cohorts, but neither is significant. Thus, while the subsample results show a slight tendency for boys to be hurt more, there are no clear differential effects by parental education and grade level.

Table 5: Regression results. Heterogeneous effects.

	Main effect: exposure in reform years		Interaction: difference, normal years		
	Coef.	*se*	*Coef.*	*se*	*N*
By gender					
Male	-0.0430(*)	(0.0245)	0.0248	(0.0339)	122,512
Female	-0.0318	(0.0219)	-0.0066	(0.0313)	124,244
By maternal education					
No tertiary education	-0.0362(*)	(0.0215)	0.0142	(0.0308)	156,994
Tertiary education	-0.0396	(0.0270)	-0.0023	(0.0366)	89,762
By grade level					
Exposed in grades 5-6	-0.0322	(0.0293)	0.0074	(0.0391)	131,643
Exposed in grades 7-8	-0.0408	(0.0268)	0.0068	(0.0394)	115,113

Standard errors are reported in parentheses and clustered by grade/school/school-year cell.
** p<0.05, ** p<0.01, *** p<0.001*

Mechanisms

The results of the analysis suggest that the only detectable negative effects of exposure are driven by the policy induced flows of SEN students during the reform years of 2012/13 and 2013/14. In this section, I explore potential mechanisms. First, the

[39] Parental education is measured as the mother having (some) tertiary education or not.
[40] This is done by estimating separate estimations for the two cohorts: students in the younger cohort who contribute with gain scores from grades 2-4 & 4-6 and those in the older cohort who contribute with gain scores from grades 4-6 & 6-8.

additional returners may either be spread across more schools, or each school may receive more returners than usual (or both). When returners are spread across more schools, schools that are not used to receive returners may face the (for them) new and complex task to accomodate learners in regular classrooms who previously have been taught in segregated settings. These schools may not have the same skills and experience to cope with the situation as schools that are used to receive returners, and this might affect learning in regular classrooms. Calculations show that while 60% of public schools receive returners in non-reform years, in reform years this share is 70%. Thus, a number of schools that perhaps usually not receive returners[41], do so during the reform years. If these schools do less well in accomodating returners in regular classrooms, this may be part of an explanation of the negative externalities of returners I find for the reform years.

However, during the reform years, the returners may not only be spread across more schools, but receiving schools may on average take in higher numbers of returners than in normal years. This would increase treatment intensity because more students would have returners as class mates compared to just grade mates, or there might be more than one returner in more classrooms. Calculating the average number of returners per receiving school in reform years and normal years, the number of returners per school in reform years is 20% higher than in normal years. Thus, this could also contribute to explain the negative effect of exposure during reform years.

Another mechanism that may be at play is if the additional returners who are brought back as part of the policy effort differ from the average returner in non-reform years. For example, if they have more serious difficulties than the usual returners this may also be part of an explanation of the negative effect for the reform years. We have no direct measure of the degree of difficulties, but as an approximation we consider the share of students who return to regular classrooms from special schools rather than from special classes in mainstream schools. Students who have been educated in special schools rather than special classes may have more serious difficulties on average - both because they have been sent to a specialist school which may indicate a higher degree of special needs, but also because returning from a separated school may be a bigger step than returning from a separated classroom within the mainstream school. Yet, I find almost no difference in the share of returners from special school and special class across reform years and normal years.[42] However, this measure is only a vague approximation to the degree of difficulties students may have. Therefore, no strict conclusions should be drawn on this piece of evidence alone.

[41] However, one cannot rule out that these schools previously have received returners.

[42] In the reform years, 81% of returners are from special classes and 7% from special schools. The corresponding numbers in other years are 80% and 9%.

6. Conclusion

Not all students educated in segregated academic settings stay for their entire school career, but many eventually return to regular classrooms and must be re-integrated upon return. These return flows may result from conventional movements of SEN students between segregated and inclusive settings or from policy reform aiming to raise inclusion rates.

This study examines potential externalities of SEN students who return from segregated education settings to regular classrooms on academic achievement of their peers in the receiving school-cohorts. While previous studies analyse the effects of the presence of SEN students in schools and classes, this study explicitely focusses on SEN students who previously have been educated in separate settings. When these students return to regular classrooms, they need to adapt to a regular class environment which – at least during a transition period – might yield challenges for themselves, their teachers and peers.

The estimation strategy exploits a full population panel dataset covering six years of data with repeated observations of test score gains over time for individual students. By comparing achievement gains of peers who in one period are exposed to a recent returner, but not in another period, the repeated gain score measures are exploited to identify effects of exposure on students in receiving cohorts. Additionally, the estimation strategy includes school-by-grade fixed effects, and thus, the estimation strategy allows for any time-invariant individual characteristics of the student and for differences in school quality across grade-levels within a school. Furthermore, the analysis is enhanced by another dimension to investigate the differential effect of being exposed to returners in years with reform-induced higher return rates. This yields a triple differences identification strategy that provides a particularly strong causal identification.

The results show a small negative effect of returning SEN students on peers' reading achievement during reform years when extraordinary large numbers of student return (-0.037 SD). The effect size is equivalent to roughly 5% of the initial test score gap between students from well-educated and less well-educated homes. The effect of becoming exposed during non-reform years is not significantly different from zero. While subsample results show variation by gender (the exposure effect being slightly larger for boys), results do not vary systematically by parental education and grade level.

To conclude, this study suggests that in years with normal levels of return flows, accomodating returners in regular classes does not seem to hurt average peers' learning. However, sending back much larger than usual numbers of returners to perhaps not sufficiently prepared schools has negative externalities for peers' learning. Even though the estimated adverse effects are of modest size, this suggests that reforms aiming at

creating more inclusive school systems should be carefully designed and prepared prior to rollout.

References

Ammermueller, A. & J.-S. Pischke (2009): Peer effects in European primary schools: Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics* 27(3): 315-48.

Andersen, S.C., Beuchert-Pedersen, L.V., Nielsen, H.S. & M.K. Thomsen (2014): 2L Report. Analysis of the effect of two teachers programs. Report (In Danish). Department of economics and business economics, Aarhus University, and Trygfonden's Center for Child Research with contribution from Rambøll Management Consulting.

Angrist, J. D. & K. Lang (2004): Does school integration generate peer effects? Evidence from Boston's Metco program. *American Economic Review*. 94(5): 1613–1634.

Angrist, J. D. & J.-S. Pischke (2009): *Mostly Harmless Econometrics: An Empiricist's Companion*. Massachusetts Institute of Technology.

Carrell, Scott E., and Mark L. Hoekstra (2010): Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. *American Economic Journal: Applied Economics*, 2(1): 211-28.

Cho, R. M. (2012): Are There Peer Effects Associated with Having English Language Learner (ELL) Classmates? Evidence from the Early Childhood Longitudinal Study Kindergarten Cohort (ECLS-K) *Economics of Education Review*, 31(5): 629-643.

Cullen, J. B. (2003): The impact of fiscal incentives on student disability rates, *Journal of Public Economics.* 87(7-8):1557-89.

Diette, T. M., and R. U. Oyelere (2014): Gender and Race Heterogeneity: The Impact of Students with Limited English on Native Students' Performance. *American Economic Review*, 104(5): 412-17.

Epple, D. & R. E. Romano (2011): Peer effects in education: A survey of the theory and evidence. Handbook of Social Economics. 1B: 1053–1163.

European Agency for Development in Special Needs Education (2012): *Special Needs Education. Country data.*

Figlio, David N. (2007): Boys Named Sue: Disruptive Children and Their Peers. *Education Finance and Policy*, 2(4): 376-394.

Fletcher, Jason 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.

Fletcher, Jason M. (2010): Spillover effects of inclusion of classmates with emotional problems on test scores in early elementary school. *Journal of Policy Analysis and Management*, 29 (1): 69-83.

Friesen, J., R. Hickey & B. Krauth (2010): Disabled Peers and Academic Achievement, *Education Finance and Policy*. 5(3): 317-348.

Gottfried, M. A. (2013): The Spillover Effects of Grade Retained Classmates: Evidence from Urban Elementary Schools. *American Journal of Education*, 119:405-444.

Gottfried, M. A. (2014): The achievement effects of tardy classmates: evidence in urban elementary schools, *School Effectiveness and School Improvement*, 25(1): 3-28.

Gottfried, M. A., & A. Harven (2015): Classmates with Emotional and Behavioral Disorders and the Protective Nature of Peer Gender. *The Journal of Educational Research*, 108: 45-61.

Hanushek, E. A., J. F. Kain, & S. G. Rivkin (2002): Inferring Program Effects for Specialized Populations: Does Special Education Raise Achievement for Students with Disabilities. *Review of Economics and Statistics*. 84(4): 584-599.

Hoxby, C. M. & G. Weingarth (2007): Taking race out of the equation: School reassignment and the structure of peer effects. *Education*: 18, 2005.

Jacob, B. (2005): Accountability, incentives and behavior: evidence from school reform in Chicago. *Journal of Public Economics*, 89: 761-96.

Kristoffersen, J. H. G.; Krægpøth, M. V.; Nielsen, H. S.; Simonsen, M. (2015): Disruptive School Peers and Student Outcomes. *Economics of Education Review*, 45(4):1-13.

Kwak, S. (2010): The Impact of Intergovernmental Incentives on Student Disability Rates. *Public Finance Review*, 38(1): 41-73.

Lavy, V., Paserman, M. D. & A. Schlosser (2012): Inside the black box of ability peer effects: Evidence from variation in the proportion of low achievers in the classroom. *Economic Journal*. 122(559): 208–237.

Lavy, V., Silva, O. & F. Weinhardt (2009): The good, the bad and the average: Evidence on the scale and nature of ability peer effects in schools. NBER Working paper 15600.

Lazear, E. (2001): Educational production. Quarterly Journal of Economics. 116(3): 777–803.

Lefgren, L. (2004): Educational Peer Effects and the Chicago Public Schools. *Journal of Urban Economics*. 56(2): 169-191.

McEwan. P. J. (2003): Peer effects on student achievement: evidence from Chile. *Economics of Education Review*. 22:131–141.

Ministry of Education (2015): Special needs education and inclusion. (In Danish: Specialundervisning og inclusion). Report. Copenhagen.

Ministry of Finance (2010): Special needs education in primary and lower secondary school – towards better organization and management (In Danish: Specialundervisning i folkeskolen – veje til en bedre organisering og styring). Report. Copenhagen.

Rangvid, B. S. & M. Lynggaard (2014): Special needs students' results at the end of compulsory education and 5 years later. (In Danish: Specialskoleelevers resultater ved skolegangens afslutning og 5 år senere). Danish School of Education, Aarhus University.

Sacerdote, B. (2011): Peer effects in education: How might they work, how big are they and how much do we know thus far? *Handbook of the Economics of Education*. 3: 249–277.

Todd, E. P. & Wolpin, K. I. (2003): On the specification and estimation of the production function for cognitive achievement. Economic Journal, 113:F3–F33.

World Health Organisation (2011): World report on disability. Report. Geneva, Switzerland.