Jon Marius Vaag Iversen

Educational Governance

Interactions Between Institutions and School Resources

Thesis for the degree of Philosophiae Doctor

Trondheim, April 2013

Norwegian University of Science and Technology
Faculty of Social Sciences and Technology Management
Department of Economics

NTNU – Trondheim
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Social Sciences and Technology Management
Department of Economics

ISBN 978-82-471-4216-5 (printed ver.)
ISBN 978-82-471-4217-2 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2013:60

Printed by NTNU-trykk

Contents

Preface

Acknowledgement

1 Introduction and Summary

2 Disadvantaged students in the early grades: Will smaller classes help them

3 School Accountability Reforms and The Use of Special Education

4 Conditional Peer Effects: The Role of School Authorities

5 Are Non-Eligible Students Affected by Special Education

Preface

This thesis consists of an introductory chapter and four independent essays on topics within the economics of education. The essay in chapter 2 is joint work with my supervisor Hans Bonesrønning (Norwegian University of Science and Technology), and the essay in chapter 5 is joint work with my supervisor Hans Bonesrønning and Ivar Pettersen (both Norwegian University of Science and Technology). The essay in chapter 2, entitled Disadvantaged students in the early grades: Will smaller classes help them?, is forthcoming in *Education Economics.*

Acknowledgements

This thesis is written as a part of the project “Governance, management and performance in the Norwegian education system”, financed by the Norwegian Research Council. It is a result of the effort from a long list of people. Their suggestions and contribution has been of invaluable importance to me. I am especially grateful for the contribution from my supervisor Hans Bonesrønning. His guidance and support through the last years have been of high importance. Hans is also a co-author on two of my papers. Hans guided me through my master thesis some years ago and introduced me to the research in education economics. His feedback and his suggestions are always very constructive and well founded and he has always found available time for discussions. I would like to thank Ivar Pettersen, who is my second co-author in chapter 5. My co-supervisor has been Lars Erik Borge. His comments on chapter 4 were important in completing this paper.

Through these years I have been part of a discussion group at Department of Economics at NTNU together with professors and phd students within the field of education economics. I am very grateful to this group for very constructive discussions on my papers as well on other topics within the economics of education. The participants of this group are Torberg Falch, Bjarne Strøm, Arnt Ove Hopland, Marianne Haraldsvik, Ole Henning Nyhus, Kaja Høiseth Brugård, Rune Reiling and Lars Erik Borge. I am also grateful to other occasionally participating in the group.

In 2012 I had the privilege to visit Lancaster University for a 7 months long stay. I am grateful to Professor Steve Bradley for inviting me to the Department of Economics at the university, and the possibility of being a part of the research group at Lancaster University. A

special mention goes to Colin Green, who was very helpful with many of the practical matters, both before and during our stay. Colin has also read and commented on two of my papers. Both chapter 3 and chapter 4 were improved by his comments.

I appreciate comments on early drafts of each chapter. Two anonymous referees deserve thanks for many suggestions in chapter 2. Also participants at the department seminar at NTNU, the 4th conference of the RTN Network "Economics of Education and Education Policy in Europe", and the workshop on "Labor and Education" at Refsnes, Norway.

Chapter 3 has benefitted for comments at a department seminar in Lancaster Univeristy, a department seminar at NTNU Trondheim, the workshop on Educational Governance and Finance in Oslo, the annual meeting in Scottish Economic Society in Perth Scotland and the third international workshop on Applied Economics of Education in Catanzaro, Italy. For chapter 3 I am also very grateful for comments from David Figlio on an early draft. His comments were important in my choice of empirical strategy.

I have received important comments on paper 4 from five anonymous referees. These comments have led to substantial improvement on an early draft. I thank the participants at the 33th annual meeting of the Norwegian Association of Economists, the workshop on Educational Governance in Trondheim and the 24th annual EALE conference in Bonn, Germany.

Chapter 5 has benefited from comments from participants at the area conference on the economics of education at CESifo. I am also grateful for comments from several of the researchers involved in the project "Education 2011" organized by Statistics Norway.

Finally, and importantly, I would like to thank my family: My two sons, Filip and William. You remind me of what is really important in life, to Ann Lisbeth for all your love, support and patience.

1. Introduction and summary

Are governing systems important in educational production? There is a heated discussion about the merits of various systems among politicians and stakeholders, and the existing empirical research is not conclusive. This thesis contributes by providing empirical evidence about the workings of governing systems in the Norwegian education system.

Up until 2006 the Norwegian governing system was characterized by strict national rules and regulations that strongly influenced the number of students per teacher, teacher workload and teacher qualifications. Norwegian investments in education were, and still are, substantial. Norway invests about USD 15000 per student, compared to an OECD-average on USD 7600. Only Luxembourg has higher spending among OECD countries. About 6,3 % of GDP are spent on education in Norway, compared to 4,6 % in OECD countries (OECD, 2012).

It is fair to say that the majority of politicians believed that this was a well-functioning system, and thus that the Norwegian students' performance on international tests from 2000 and onwards was far below expectations (PISA, TIMSS). In addition, the findings reported by Hanushek (1986) that there is no systematic relationship between class size and student achievement were important when policy makers started to rethink the governing system.

The combination of high expenditures, mediocre student performance and existing empirical evidence on class size effects motivated the reform in 2006. "Kunnskapsløftet" (Knowledge promotion) was introduced at the national level in Norway, and was a very broad reform with many aspects. However, important parts of the reform contained governance and management. In a report to the national parliament these encouragements was described

(Kunnskapsdepartementet 2004) The new governing systems should be based on basic principles about:

- Clear national goals
- Information and knowledge about performance and results
- Clarifications of responsibility
- Local freedom to make decisions
- Guidance and support

The understanding that the governments could create equal educational supply through strict national regulations and governance is replaced by the reliance that teachers, school principals and school owners have the best opportunities to produce educational outcomes within the national goals. The schools were meant to be freer to do adjustments based on the needs and qualities of the student group.

From a theoretical point of view, the effects of increased spending or reduced class size may be conditional on the prevailing institutional and government settings. Governing systems based on strict regulations such as the student-teacher ratios do not impose incentives to improve performance. One of the main goals with modern accountability systems is to impose important incentives for school principals, teachers and students to improve academic achievement within the budget constraint. Therefore, investigations of causal impacts from school factors on student performance may act differently under various governing systems. In this thesis I investigate the use and impact of school resources in different governing systems, both before and after the implementation on “Kunnskapsløftet” in 2006.

First, I study whether reduced class size is effective in the traditional governing system, with focus on disadvantaged students. These analyses are inspired by Angrist and Lavy (1999).

Second, I go on by focusing on the use of special education, which has increased substantially after 2006. These analyses take place in modern governing systems characterized by decentralization and to some extent accountability. Accountability systems are often motivated by productivity goals. There are higher demands for student achievement given a budget constraint. However, worldwide we have seen situations where reform implementation has caused higher use of resources, especially through the use of special education (See Jacob 2005, Figlio and Getzler 2006 and Cullen and Reback 2006). In this thesis, I focus on the reasons for the increase in the proportion of students that are special education placements in Norway, and its consequences.

In chapter 3 I investigate reasons for the increase in special education resources, especially whether this increase differs between reformed and non-reformed municipalities. Because the Norwegian institutions are characterized as a federal system, I have to distinguish between the national part of the reform and the municipal part of the reform. All municipalities have to take into account the national part of the reform. The most important part at the national level is the introduction of national tests. There is evidence that municipalities have implemented the reform to varying degrees. Some municipalities had elements of accountability prior to the reform, while many municipalities have implemented such elements in the years after the national implementation. There are municipalities that still not have implemented any accountability elements in their governing systems. At the same time, annually national tests for 5^{th} and 8^{th} graders were implemented from 2007. There were also national tests in 2004 and 2005. In chapter 3, I distinguish between non-reformed and reformed municipalities based on their level of reform implementation at the municipality level, and analyze whether these groups of municipalities experience differences in the increase of special education.

In chapter 4, I do also focus on the use of special education. This chapter has to aims. First, I present evidence that peer effects cannot be estimated without taking the use of special education into account. Second, I investigate whether the use of special education has an impact on the estimated peer effects. In chapter 5, I study the effects of special education resources further. I investigate whether special education placements and hours of special education per student has an impact on the student's peers. Are the other students in the class affected by the decision of giving a student special education?

This thesis contributes both to the literature on governing systems, and to the literature on the effects of special education. The thesis have policy implications, especially for the school owners. To achieve the effects of the national part of the governing reform, municipalities have to respond by implementing accountability systems. If not, they will not be able to handle the increased demand for special education, coming form the national aspects of the reform. This thesis does give implications for the use of special education. There seem to be positive effects of such school resources, but it depends on the organization. By increasing the number of special education hours per student, we might see a positive effect on the students' peers.

I will now give a brief summary of my papers.

Chapter 2:

Disadvantaged students in the early grades: Will smaller classes help them?

The aim for this paper is to analyze the traditional governing system in Norwegian educational system. It was characterized by strict national regulations, and the quantifying of school inputs. One such kind of national regulations was the class size rule saying that if

enrollment were above 28 students, the group should be divided into two classes. Similarly, if enrollment exceeded 56 students the group should be divided into three classes. This rule was abolished in 2003-2004. In this paper we exploit the class size rule, following the regression discontinuity (RD) approach - first introduced in the economics of education by Angrist and Lavy (1999), and recently applied on Norwegian data by Leuven, Oosterbeek and Rønning (2008).

The first important feature of this analysis is that we can credibly investigate a cumulative class size effect on student performance for students that are 10 years old when they are tested. The second feature of this analysis is that it focuses on disadvantaged students. Two approaches are used towards this end. In the first part of the paper we investigate class size effects for students that are more or less disadvantaged using the conventional education production function framework. Thereafter we investigate whether the slope of the relationship between student achievement and family background characteristics are conditional upon class size. This latter approach is complementary to the conventional education production function approach and focuses more directly on the potential of class size reductions as a remedy for reducing socioeconomic differences in student achievement, or alternatively, this analysis sheds light on the potential heterogeneity in class size effects.

Policy implications from this paper are that while class size reductions are not effective on raising students' average performance, it may be used as a tool to reduce socioeconomic differences, when improving educational achievement for disadvantaged students.

Chapter 3:

School accountability reforms and the use of special education

This paper analyzes the effects of modern governing systems in Norway, following the national reform "Kunnskapsløftet". There is a large international literature on such effects. Due to methodological issues, however, credible causal effects are difficult to obtain. We can distinguish between intended and unintended effects of accountability reforms. My paper focuses on the unintended effects, more specifically on the impact of reform implementation on the use of special education.

The level of special education has increased rapidly since the implementation of "Kunnskapsløftet". However, the municipal implementation of the governance elements in "Kunnskapsløftet" has been done to a varying degree. The aim of this paper is to analyze the relationship between reform implementation and the use of special education.

I exploit the variation in time of implementation, when using a generalized difference in difference model, introduced by Jacobsen, Lalonde and Sullivan (1993). This model is useful, in distinguishing between long term and short term effects, early and late implementers, and the model is exploiting the variation in time of implementation

The results in this paper indicate that the increase in special education is significantly lower in municipalities that have implemented the reform. I argue that this is because municipalities are better equipped to keep special education down in such a governing system. There is more information available, such that the information asymmetry between the school owner and the teacher is lower. Special education is an expensive type of teaching and the school owner wants to keep this level down in order to allocate more resources to the majority of students.

After introducing accountability system the school owner are more able to keep special education at a lower level.

Chapter 4:

Conditional peer effects and the role of school authorities.

The literature on peer effects in educational production is extensive. However, little is known about the sign of the effects, because of methodological issues. The current paper deviates from the existing gender peer group literature in several important ways. Most importantly, I emphasize the role of special education in estimating peer effects, as the relationship between classroom gender composition and student achievement seems to be conditional on the level of special education. Through these exercises, I indirectly identify a potentially important tool, in which politicians, principals and administrators are able to attenuate negative externalities in the classroom.

Norwegian institutions are characterized by full inclusion and extensive use of special education. In this setting, it might be hard to estimate credible peer effects, without taking account of the use of special education. I present evidence that school principals and school owners use special education to reduce the negative externalities of bad peers. Using special education resources to disruptive students will reduce the negative gender peer effects, and we have to take this into account when interpreting peer effects as causal.

My paper has two parts. First, I follow the existing literature in detail when estimating school fixed effects models. This has been a common strategy for this kind of analyses since Hoxby (2000a). An important contribution is Lavy and Schlosser (2011). Inspired by the theory of Lazear (2001) they highlight negative externalities in educational production, and they find

that disruption is one of the main mechanism in which the gender peer effects operate. Initially I follow their approach as a motivation for my analyses, and find negative gender peer effects as well. However, I find a strong relationship between the level of special education and the gender composition. The peer effects seem conditional upon the level of special education.

The second part of the paper seeks to dig more deeply into the relationship between peer effects and special education. I use different strategies. First, I try to estimate the peer effects surroundings when I don't need to take account of special education. Second, I will through this exercise investigate the magnitude of special education on the educational setting. Exploiting variations in special education over time I compare the same schools at different points of time based on their level of special education. My findings indicate that school authorities use special education as a tool to improve the learning environment.. The use of special education has a strong effect on the educational setting, in reducing negative externalities from bad peers.

Chapter 5:

Are Non-eligible Students Affected by Special Education

In chapter 4, I identify special education as an important policy tool, when affecting the educational setting in the classroom. The analyses in chapter 4, does not answer any questions of the effect of such activity on student performance. Having the results in chapter 4 in mind, I analyze whether the performance of non-eligible students are affected. Special education is specially designed instruction that aims at improving the performance of students, who because of a disability, do not benefit from ordinary classroom teaching. The purpose of the

present paper is to investigate whether such resources have non-intentional effects by benefitting non-eligible students.

Using data from the Norwegian elementary school, we start out by providing evidence that the academic performance of non-eligible students is negatively correlated with the proportion of students that are eligible to special education, presumably because misbehaving students are overrepresented within this subgroup. To investigate whether special education resources dampen the negative externalities, we take advantage of a large, across-the-board increase in the proportion of eligible students, and combine school fixed effects with an IV-approach to identify the causal effects of special education on the academic performance of non-eligible students. We find that non-eligible students are positively affected by *the number of hours in special education per eligible student*, while there seems to be negative effects from increasing *the proportion of eligible students*. Thus, our estimates indicate that non-eligible students experience substantial achievement gains when schools, for a given special education budget, choose to allocate more of these resources to a few students instead of providing many students with less special education resources.

Discussing chapter 3, chapter 4 and chapter 5 together, we have broad evidence on the effects of special education. Chapter 3 presents evidence that municipalities that have implemented accountability have a significantly lower increase in special education than non-reformed municipalities. In chapter 3 I find that the gender peer effects are conditional upon the level of special education, and I find indications that the use of special education has a positive impact on educational setting. These mechanisms are important for interpreting the results in chapter 5. The improvement in the educational setting from more special education per student results in an improvement in performance for non-eligible students.

Further research

All my four papers in this thesis motivate further research. Chapter 1 answer questions about the effects of class size for disadvantaged students, but does not answer questions about mechanisms in which the class size operate. The way teachers react to different class size might be of importance. However, missing data about Norwegian teachers has so far made it impossible to highlight these hypotheses.

Just a few papers analyses reform implementation of Norwegian education governance. I have presented evidence that municipalities which have implemented accountability elements experience less increase in the level of special education. Further investigations should answer questions about the intended effects of reform implementation. However, existing data has not sufficient common time variation in performance data and reform implementation. We are in need of a credible instrumental variable strategy to complete these analyses. Alternatively, one interesting research question is to dig more deeply into the relationship between reform implementation and resources. Why are reformed municipalities increasing their level of special education less than non-reformed municipalities? Possible answers to that question is that they allocate more resources to the rest of the students, that they are using less resources overall, both to students with special needs and other students, or that the give more special education to students. This hypothesis could be investigated further.

My peer effects analyses could have been extended in several ways. It is possible to highlight the heterogeneity of peer effects, or more importantly, follow Lavy and Schlosser (2011) in order to identify mechanisms in which the peer effects operate, or mechanisms in the relationship between special education and the estimated peer effects. Data on the educational

setting and learning environment exists only on school level. However, for this research questions, school level analyses could have answered some of the unsolved issues.

Chapter 4 and chapter 5 have much in common. Future research should investigate the relationship between the issues raised in these two chapters into more detail. Chapter 4 indicates that the possible mechanisms in which student performance is improved, is from the effect of special education on the educational setting and learning environment. The negative externalities from bad peers are reduced. These issues could be investigated further, through analyzing data on learning environment as dependent variable. Does the level of special education affect these measures for student satisfaction? Both these data sources are available today at the school level.

Paper I

Is not included due to copyright

Paper II

SCHOOL ACCOUNTABILITY REFORMS AND THE USE OF SPECIAL EDUCATION

Jon Marius Vaag Iversen

Centre for economic research and Department of Economics, NTNU

N-7491 Trondheim, Norway

Abstract

This paper analyses the relationship between Norwegian school reform implementation and the use of special education. After the introduction of the national reform programme called the "Kunnskapsløftet" ("Knowledge promotion") in 2006, the use of special education has increased dramatically in Norway. As part of the national reform, municipalities were encouraged to implement accountability elements in their governing systems. There is evidence that the municipalities have implemented the reform to varying degrees and at different points in time. I exploit these variations in timing to investigate whether the growth in the use of special education reflects the degree of reform implementation in these municipalities. The variation in the timing of reform implementation is exploited by means of the application of a Jacobsen, Lalonde, and Sullivan (1993) strategy. I find that increases in the proportion of special education placements are significantly smaller in municipalities with a high degree of reform implementation.

Keywords: School accountability, special education

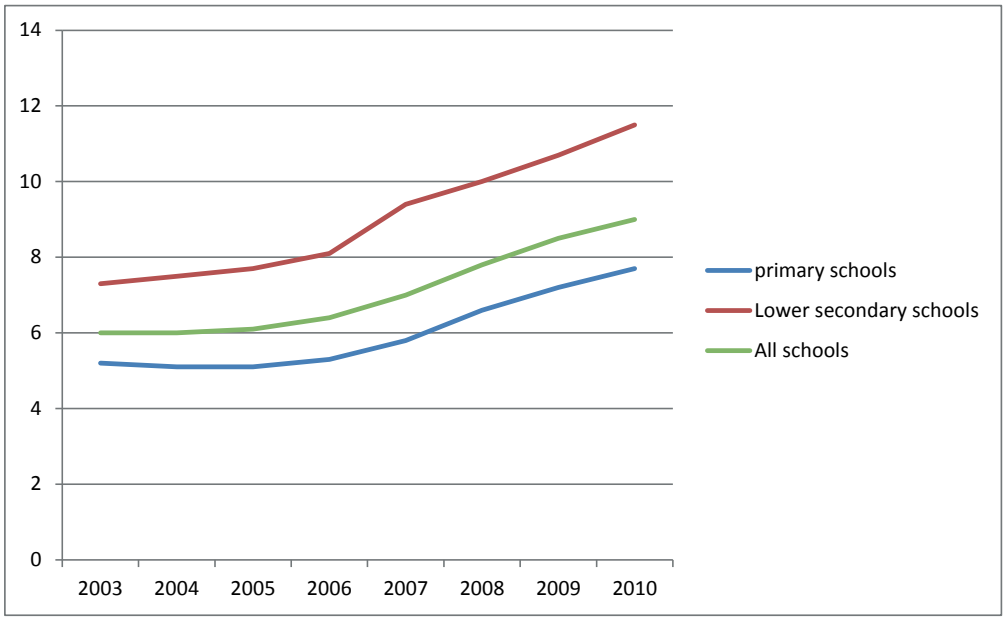
*E-mail: jon.iversen@svt.ntnu.no

1. Introduction

Evidence from the existing international literature on the effects of accountability in education is mixed. With respect to intended effects, several studies suggest that high-stakes testing may be effective at raising student achievement (i.e., Carnoy and Loeb 2002, Jacob 2005, Dee and Jacob 2011 and Hanushek and Raymond 2005), but Ladd (2009) finds small achievement benefits from the implementation of accountability policies in Dallas in the 1990s. Estimating unintended effects, several authors have highlighted aspects of strategic responses from schools. For example, there is evidence that schools respond to accountability pressures by differentially reclassifying low-achieving students as learning disabled so that their scores will not count against the school in the accountability system (See Cullen and Reback (2006), Figlio and Getzler (2006), Jacob (2005)).

In Norway, a national educational reform act that involved elements such as testing, decentralisation and accountability was passed in parliament in 2006. Since that time, the country has witnessed a significant expansion in the number of special education placements in elementary schools (see figure 1). The proportion of special education placements was fairly stable at approximately 6 percent until 2005-2006, when it began to increase, reaching 9.1 percent by 2010; this change represented an increase of more than 50 percent. This growth trend is similar for primary and lower secondary schools.

FIGURE 1: The growth of special education in Norwegian schools since 2003.



There is heated debate in Norway about increasing resources for special education. Some of the participants in this debate think that Norwegian schools have responded to accountability in the same way as schools elsewhere, i.e., by strategically reclassifying low-achieving students as learning disabled. The hypothesis that the increase in special education placements is caused by the reform cannot easily be investigated because of the lack of a counterfactual (control group). In this paper, however, I take advantage of the fact that Norway has a federal governing system in which municipalities decide whether to implement reforms. Parliament has passed the national reform, introduced the national tests, and initiated the publication of test results at the municipality level. However, unless the municipalities that oversee the schools implement the reform by introducing new governing systems in their own districts, school leaders will experience no more accountability than they did prior to the reform.

Nevertheless, all the main actors in education in all the municipalities are influenced by the new informational environment that was introduced by the national tests.

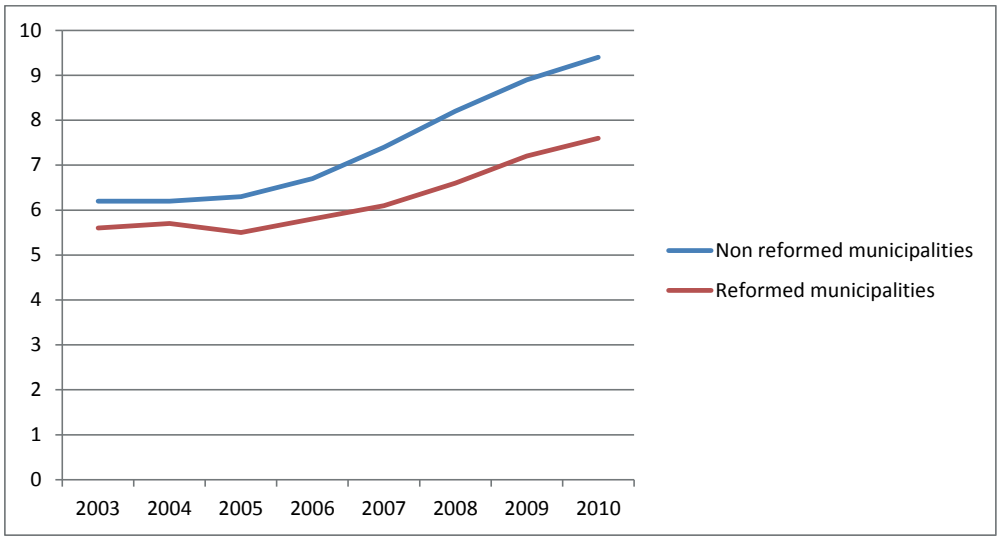
My argument proceeds along the following path. Because of the national aspect to the reform, the group of parent-customers will be better informed and is likely to put more pressure on schools. In particular, there might be increased demand for more resources for low-achieving students. Teachers may respond in a similar fashion, by demanding that more students be reclassified as learning disabled. Thus, it seems likely that the national part of the accountability system will increase the demand for special education placements by teachers and parents. Conversely, the local county council is elected by inhabitants of the municipality. Some groups of voters, such as teachers and other public employees, are reluctant to implement the reform. The relative strength between groups of voters will decide the composition of the local council. If the local county council decides to introduce a new governing system, it will be better equipped to keep costs and special education placements down by accessing information about student qualities and holding school leaders accountable.

My empirical strategy is to exploit the variations in the timing of reform implementation across municipalities to investigate whether properly implemented reform causes more moderate growth in the share of special education students. I obtained much of my information about the implementation of the national reform from surveys to the chief executives in the municipalities. The information contained in the survey indicates that as of 2009, as many as two-thirds of the municipalities had not yet introduced the accountability elements contained in the reform into their systems. The survey data are combined with information about special education placements from a national database. In Figure 2, I show

the development of special education placements for the implementing and non-implementing municipalities separately. Figure 2 shows that the reform-implementing municipalities have experienced less growth in the share of special education students compared to the non-implementing municipalities. These findings are the first indications about the relationship between municipality reform implementation in Norway and the use of special education. However, more sophisticated strategies are needed to determine causation.

In the empirical analyses, I include an application of the approach developed by Jacobsen, LaLonde & Sullivan (1993) (JLS) to exploit variations in the timing of reform implementations in Norwegian municipalities. I compare the within-municipality proportion of special education placements in the years after the reform with the proportion of special education placements for a point of time many years prior to the reform; in the next step, I compare the within-municipality variation to the national trends. With this strategy, I can distinguish between the short- and long-term effects. The results from these analyses confirm the information represented by figure 2, i.e., that non-reformed municipalities have more significant growth in the share of students placed into special education than reformed municipalities.

Figure 2: Growth in special education for reformed and non-reformed municipalities



The rest of the paper is organised as follows. Section 2 presents Norwegian institutional settings. These settings are discussed in the context of a three-tier principal-agent framework that guides the empirical investigations in this paper. I also present the existing empirical literature in this section. Section 3 presents the data and the empirical strategy. Section 4 provides results and section 5 concludes.

2. Theoretical considerations and institutional settings

The choice of the theoretical framework for this analysis is motivated by the following stylised argument. The various actors within the educational system might have responded in different ways to the reform and the information provided by national tests. First and foremost, the available evidence seems to indicate that there is substantial variation in the school owners' responses; some have installed the accountability systems and others have not. I have no information regarding the responses of the teachers and parents. However, if they

have responded uniformly by demanding more special education resources for low-achieving students, then all school owners have experienced an increased demand for special education resources. The owners' capacity or willingness to sustain lower costs might depend on the governing system, i.e., the increase in special education placements might be more moderate in municipalities where the leaders are held accountable for school productivity.

These theoretical considerations should in this case be discussed in conjunction with a discussion of the Norwegian institutional settings in this case.

2.1 Institutional settings

Norway has a federal governance system. Local councils decide the number of schools, their locations and their budgets. In addition, the local councils decide the local governing systems, such as the degree of decentralisation of the decision-making authority and the degree of accountability. The national government imposes constraints on municipalities through various mechanisms. The municipalities' revenues are mostly determined by grants and local taxes that are decided by the national government.

In the years before the "Kunnskapsløftet" reform, the national governance system was characterised by many strict national rules and regulations that influenced the operations of the educational sector. These regulations governed class size, teacher workloads, national exams and curriculum. There was little freedom for the municipalities and schools to adjust their resource allocations.

During the time since the enactment of "Kunnskapsløftet" and in the years immediately before it was passed in parliament, many of these regulations were abolished. The reform included a change at the national level and encouragement for municipalities to change their

governing systems. This change involved both decentralisation and accountability reform. There was decentralisation in education both from the national level to the municipality level and then from the municipalities to the school level; the municipalities were encouraged to further decentralise educational decisions to the school level. The first segment of the decentralisation reform (from national to municipal government) was executed mainly by abolishing national regulations. From 2004 onwards, there was no longer a class size rule, and in addition, the first national tests were performed in 2004 and for every year that followed, except for 2006. The test results were disclosed publicly at the municipality level, and newspapers provided rankings at the school level. One important implication of publicising the results from these national tests is that it provides important information for all parts of the system. Exams, tests and curriculum are still decided at the national level.

The municipalities were encouraged to form their educational governance systems such that school principals were held accountable for student performance. The national tests were supposed to be an important reform ingredient for the benefit of the school owner in negotiating school leader contracts or in performing systematic evaluations. The test results were also to be used by the school principal to inform school-level decision making. Additionally, the national tests were to provide teachers and parents with a plethora of useful information about their students and children. Based on the information gleaned from the results of the national test, parents might pressure schools to improve their performance and they also might acquire more resources to help their own children to improve performance. There is evidence that the municipalities have implemented the reforms to varying degrees at different time intervals. This variation in the time of implementation is important in my empirical strategy presented below.

All students have the right to obtain special education if they do not benefit from ordinary instruction. The decision to give special education is the school owners' responsibility, in theory. In many cases, however, these decisions are made by the school principal. In fact, more than 70 % of the municipalities report that they decide their own level of special education. Among municipalities that have implemented accountability systems, the share that decides their own level of special education is 85%.

To make decisions about special education placements, the principal needs an expert evaluation from the educational psychological service, which is usually located in the municipality administration. The experts give advice to the principal about special education regarding whether students would benefit from ordinary instruction based on individual student evaluations. In the next step, the principal might accept or deny these recommendations; he may then decide whether to offer the student special education.

2.1 Theoretical considerations

Because the reform is national in scope, there is more information available for all parts. However, because it is obvious that teachers and parents have more information about particular student qualities, this relationship is best understood under a principal-agent framework with information asymmetry.

Drawing on early contributions by Tirole (1986) and Laffont and Tirole (1993), Dal Bo (2006) presents a three-tier principal-agent model that he applies to the relationship between the government and private firms. The government – the principal in the model – hires a regulator who specialises in learning about the industry and who may find out the true costs of the firm. The firm will then have incentives to bribe the regulator into not telling the

government when costs are low. If the regulator says he has learned nothing, the best possible contract that the government can offer the firm is one that permits the firm to determine its own rents. Conversely, the government has an incentive to offer the regulator a contract that induces him to be truthful and, simultaneously, to offer a contract to the firm that reduces its incentives for collusion with the regulator.

This framework should be transferable to the decision-making of selecting students into special education. The essential contribution of such a framework is that it allows for a third party in the principal-agent model. The regulator in my model is the school leader. The school leader's position will depend on the contract the principal offers him. The principal in this model is the school owner – the municipality – and the agents are the teachers. In addition, the teachers are influenced by the parents, who will demand extra resources.

One essential part of my model that differs from Dal Bo (2006) will be that the principal is elected by the inhabitants in the municipality every four years. A large part of the inhabitants are teachers and other public employees. The teachers might find some aspects of the reform – especially the accountability aspect – unattractive. They occupy a position that may enable them to hinder the implementation of the reform; as *swing voters* that have high stakes in the policy outcomes, they might disproportionally influence the school owner. Moreover, the teachers have allies. Other local public employees might join the teachers, either out of sympathy or because they are witnessing the introduction of disciplinary devices in their own fields (Bonesrønning 2013). A broad reform implementation literature provides evidence that public employees resist market-oriented reforms. The teachers and their allies in the public sector might therefore influence the political composition in the local council, and the relative

strength of voter coalitions will decide the willingness to implement reforms (See Christoffersen & Paldam (2003) and Rattsø & Sørensen (2004)).

One important implication of the national reform is that the information contained in the national tests is available for all parts of the educational system. In this model – and according to the literature – the additional information will increase the demand for special education from the agents (see Jacob 2005 and Figlio & Getzler 2006).

The outcome will depend on the role of the regulator. In an accountability system, the principal has leader contracts with the regulator and systematic evaluations of all schools. In such a system, the regulator's incentives for collusion with the principal are smaller. For non-reformed municipalities, the non-contracted regulator might tend to sympathise with the teachers and not provide more information to the principal. The regulator is not held accountable; in such a system, the principal is not able to use the increased information that is made available and the principal does not have access to all information in the system. Essentially, because of the national aspect of the reform, the demand for special education will most likely increase. Through these contracts and through the evaluations, however, the school owner will gain information about the student group and the need for extra resources. If teachers acquire more resources than needed, the school owner will then have more power (information) to counterbalance them. The information advantage of the teachers has diminished.

The impact of the institutional settings described above should be discussed in the context of the principal-agent framework. As discussed above, the municipality administrators' willingness to implement reform will have an impact on the degree of information asymmetry in the system. Because the national part of the reform increases the available information and

the demand for special education from parents and teachers, the municipalities that do not follow the national prompting to implement new municipal governing systems will not have the right tools to handle this increased demand. These municipalities should experience greater increases in the special education level. However, if the municipality has adapted its governmental system, the theoretical framework predicts a less rapid increase in the special education level. I have also discussed reasons why certain municipalities may or may not have implemented changes through the election process. I will use these predictions when setting up the empirical specifications.

2.3 Literature

As discussed in the introduction, there is substantial literature on the effects of accountability systems, mostly from the United States. We can distinguish between studies of intended and unintended effects of implementing such reforms. One example from the latter branch of the literature is that school principals and teachers seem to reclassify students as learning disabled to take them out of the testing pool. Figlio and Getzler (2006) find that, following the introduction of the FCAT testing programme, low-performing students from humble socioeconomic backgrounds were significantly and substantively more likely to be reclassified into disability categories that exempted them from the accountability system. In another example, Cullen and Reback (2006) examine the discontinuity in rewards in Texas's accountability system to show that schools respond to incentives to shape the test pool. Jacob (2005) also finds evidence that teachers responded strategically to the accountability policy, particularly in terms of special education placements and grade retention; nevertheless, he finds no effect on the proportion of students who participated on the standardised achievement exams.

These examples from the literature are taken from countries with different institutional settings than Norway. As Figure 1 predicts, the national reform in Norway has resulted in an increase in the special education level for all types of municipalities. However, local differences in governance systems will give the school owner different choices of mechanisms with which it may handle this increased demand. This paper studies the role of the school owner – or the principal – in the theoretical model above to use its power on the teachers and the school leaders. Through contracts with regulators, the principal will give the regulator incentives to provide information about the student group, and the principal will be better equipped to handle the increased demand for special education resources. This paper will analyse the role of the principal, paying particular attention to how different governing systems affect the relationship between principal, regulator and agent in the municipality school system.

3. Data and Empirical strategy

The data used in these analyses have several advantageous features. First and foremost, I will exploit the fact that municipalities have implemented the national reform to various degrees and at various points in time.

I use data from a number of sources. First, I use school data from the school administrative system in Norway (GSI, Grunnskolens informasjonssystem). From this system, I can identify schools and find their use of special education and assistants, group size and certain information about teachers. In addition, I have merged these data with observable information about the municipalities. These data are provided by Statistics Norway.

Two surveys of the Norwegian municipalities are used to generate information about reform implementation. Both surveys were provided by researchers at the Centre for Economic Research at NTNU (Strøm et al, 2009). School owners are asked to characterise their governing systems in terms of whether they are decentralised and whether – or how – school principals are held accountable for student performance. In total, 297 of the 434 municipalities responded to the first questionnaire. A randomly chosen sample of 117 municipalities responded to the final questionnaire.

The dependent variable in my main analyses is created as the proportion of students at the school who have special education placements, i.e., the number of students receiving special education divided by the number of students at the school. I estimate regression models with the proportion of students who have special education placements as the dependent variable, and I use the initial level of special education placements as a control.

To distinguish between reformed and non-reformed municipalities, I use the information about accountability systems provided by the two surveys described above. This information is captured by the following two pieces of information: whether the school principals have signed a contract with explicit student performance objectives and whether the school owner systematically evaluates the school principals. Descriptive statistics from these questions are presented in the Appendix. Approximately one-third of the municipalities have school leader contracts, and about the same proportion performs systematic evaluations. However, the share of municipalities with systematic evaluations is increasing, while the share of municipalities with leader contracts remains fairly constant.

Table 1: Descriptive statistics

Variable	All municipalities	Survey municipalities	Non-survey municipalities	High degree of reform implementation	Low degree of reform implementation
Proportion of youths (0-16)	0,2	0,2	0,19	0,2	0,2
Proportion of highly educated	0,14	0,15	0,14	0,17	0,14
Number of inhabitants	11442	13970	7219	25441	6964
Proportion of socialists in the municipal council	38,5	38,5	*	38	38,9
Herfindahl index	24,6	24,6	*	22,6	25,8
Use of assistants	19,5	19,7	19,2	19,2	20
Proportion of teachers without licenses	0,04	0,04	0,04	0,03	0,04
Student-teacher ratio	0,1	0,11	0,1	0,11	0,1

NOTE: Survey municipalities are those responding to the survey. High degree of reform implementation is municipalities with both leader contracts and systematic evaluation

Other explanatory variables included in my models are time dummies and information about the schools, such as the use of assistants, number of students, the share of male teachers and education level among teachers. Descriptive statistics from these variables and other characteristics about the municipalities are shown in table 1.

Table 1 also provides descriptive statistics for important subsamples. First, representativeness is evaluated by comparing the group of municipalities that have responded to the survey with the group of municipalities that have not. These subgroups share many of the same

characteristics, but the responding municipalities have more inhabitants, on average. I am also able to compare the municipalities with a greater-than-average degree of reform implementation with those municipalities with a lower-than-average degree of reform implementation. These municipalities differ more. The low-reform-implementation municipalities have populations that are smaller and less educated. Of the municipalities with higher-than-average reform implementation, the number of inhabitants is, on average, more than 25000. In lower-than-average reform implementation municipalities, the average size is approximately 6000. The education level is higher in the municipalities with the higher degree of reform implementation. In municipalities that are above average with respect to reform implementation, 17 % of the inhabitants have higher education, as opposed to 14 % in the municipalities that are below average with respect to reform implementation. These differences are not very large; nonetheless, they illustrate the need for more sophisticated estimation strategies.

Table 2: Reform implementation characteristics, 2009-2010

	2009	2010
Systematic evaluations	32,7	71,5
Leader contracts	34,9	38,5

3.2 Empirical strategy

The predictions from the institutional settings and the theoretical framework discussed above seem to recommend the application of a "before/after" or "difference-in-difference" framework. Certain municipalities implemented accountability elements before 2006, some

have implemented these elements after the national reform was introduced and others have not implemented any elements yet. The national tests were introduced in 2004 and the national reform in 2006. I argued above that both events might affect the demand for extra resources from teachers and parents. Figure 2 reveals a small increase in special education from 2005, which accelerated from 2006. Figure 2 illustrates, as the theoretical framework predicts, that the increase is smaller in reformed municipalities than in the non-reformed municipalities.

To exploit the variations in time of implementation, I have chosen a strategy first introduced by Jacobsen, Lalonde & Sullivan (1993) and more recently used in Couch & Placzek (2010). This strategy makes it possible to distinguish between short- and long-term effects and between early and late implementers. It may be applied to the current data because the survey by Strom et al (2009) provides information about the year in which the municipalities changed their governing system.

3.2.1 The JLS strategy

I pool all the information for schools in the years between 2003 and 2010 and introduce a series of dummy variables for the years before or after the reform implementation. Accordingly, I let D^k_{smt}=1 if, in period *t*, school *s*, in municipality *m* had implemented the reform *k* years earlier (or if *k* is negative, if municipality m implemented the reform *k* years later).

The statistical model is given in the following equation:

$$(1)\ Y_{smt} = \alpha_m + \lambda_t + X_{smt}\beta + \sum_{k \geq -7} D^k_{smt}\delta_k + \varepsilon_{smt}$$

where Y is the level of special education in school s in municipality m at time t. Further, α is the municipality-fixed effect that captures the impact of permanent differences among schools in observed and unobserved characteristics. λ is a linear time trend that captures the general time pattern of special education in the Norwegian school system. Because the change in special education placements for reformed municipalities are captured by the reform dummies, this trend will be defined by the non-reformed municipalities. I expect a positive time trend as a result of the patterns in figures 1 and 2. The variable X is a set of time-varying controls on the school level. These include the use of assistants, proportion of male teachers, average grade size, and average educational level among teachers, in addition to the average education level among parents. The variables of interests are the reform dummies. D^{k}_{smt}, where k is in the interval of [-4 years or more before, 6 years or more after], jointly represents the time of reform implementation. In particular, the effect is interpreted as the change in special education placements for reformed municipalities k years after the reform, relative to the change for non-reformed municipalities. The point estimates for the dummy variables reflect this change as an average of all municipalities in each group. The coefficient is the gap in special education placement changes between reformed and non-reformed municipalities. If this gap increases after reform implementation, it indicates that reformed municipalities have succeeded in keeping special education placements down compared to non-reformed municipalities. Following figure 2, one should expect an increase in this gap.

The JLS analyses are in many ways a generalisation of a difference-in-difference model, allowing estimation of the average effect for many municipalities at the same time and using a comparison group to estimate the changes that would have occurred in the absence of the reform. As discussed above, this strategy distinguishes between short- and long-term effects and between early and late implementers. In addition, the JLS model allows for differences

between reformed and non-reformed municipalities before implementation. It is unreasonable to think that non-reformed and reformed municipalities were identical prior to the year of implementation. One important advantage with this strategy is that it permits the monitoring of the effects prior to implementation. Nonetheless, we know that the governing systems differ more between reformed and non-reformed municipalities after the reform implementation year than they did in the preceding years. For this reason, we expect growth rates in special education placements to differ more in the years following the reform. A JLS strategy is also more robust than a regular difference-in-difference model because it uses more of the variation in the data. I do not need to condition my analyses on one change in one municipality at one point of time.

One possible concern is that the strategy does not account for the effect of the national reform in 2006. We have included a national trend for the period, but this trend will not capture the effect of the reform. I introduce a reform year dummy to address this issue. However, we also must account for national tests existing prior to the reform. The effect of increased information available in the system discussed in the theoretical framework might have existed, to some extent, prior to the national reform. I return to this matter below. I begin from a model, as follows:

(2)

$$Y_{smt} = \alpha_m + \beta_1 D_{t>2006} + \beta_2 \,\mathrm{Im}\, pl_{smt} + \beta_3 After_{smt} + \beta_3 (D_{t>2006} * \mathrm{Im}\, pl_{smt}) + \beta_4 (D_{t>2006} * After_{smt}) + \varepsilon_{smt}$$

where, in addition to municipality-fixed effects, I have included a dummy variable, $D_{t>2006}$, which is 1 for the period after national reform implementation and 0 otherwise. $Before_{smt}$ is a dummy variable that is 1 for all the years prior to municipal implementation in school *s*, municipality *m* and time *t* and is 0 otherwise. $Impl_{smt}$ is a dummy variable that is 1 for the

years of reform implementation in school *s* in municipality *m* at time *t*, i.e., the year of reform implementation and the following two years of implementation. The variable takes the value 0 for all other years. $After_{smt}$ is a dummy variable that is 1 for the years after the second year of reform implementation in school *s*, municipality *m* at time *t* and is 0 otherwise. I have also included interaction terms between the dummy for the national implementation and both the municipal implementation dummy and the dummy indication for the years after municipal implementation. The reason for summing all the years prior to implementation and after implementation is that this model is more demanding in terms of data requirements. If there are stronger effects of local reform implementation in the time after the national reform is introduced, one would expect that the interaction term should be negative, in addition to a negative effect of implementation for the period after reform. If that is the case, the effect of the municipal reform is stronger in the period after national implementation than in the preceding period. Because the national tests were introduced in 2004 and because the characterisation of the municipal reforms before and after the national implementation is approximately the same, I would be surprised to find a significant relationship. With the results of figure 1 in mind, we expect to find a positive effect of the dummy variable reflecting the period after the national reform. The increase in this period is a break with the national trend.

Models (1) and (2) are school-level analyses. The reasons for this are that the decision to give special education in most cases is made by the school principal. Different school principals might respond differently to the same municipal governance system. However, I am aware that the decision to implement reforms is made at the municipality level. Municipality level analyses are presented as a robustness check.

4. Results

As discussed above, the data structure and the theoretical considerations indicate the value of before/after types of analyses. The JLS strategy generalises and extends traditional difference-in-difference analyses, by including a set of dummies indicating the number of years before or after reform. By using this model, I can distinguish between short- and long-term effects and can analyse whether reforms of municipal governing systems implemented after 2006 have a different effect on the special education increase than reforms implemented at an earlier stage. The JLS strategy also controls for initial differences in special education placements. This benefit is discussed in detail above, and the results are shown in table 3.

Table 3: **JLS strategy of short and long run effects of reform implementation on the level of special education**

	Level of special education	
4 years or more before reform implementation	-0.00875**	-0.00609*
	(0.00357)	(0.00356)
3 years before reform implementation	-0.0107**	-0.00686
	(0.00435)	(0.00437)
2 years before reform implementation	-0.0109**	-0.00879*
	(0.00486)	(0.00525)
1 year before reform implementation	-0.0123**	-0.0104*
	(0.00540)	(0.00558)
Reform implementation year	-0.0118*	-0.0119*
	(0.00628)	(0.00624)
1 year after reform implementation	-0.0139*	-0.0135**
	(0.00681)	(0.00675)
2 years after reform implementation	-0.0140*	-0.0122
	(0.00788)	(0.00874)
3 years after reform implementation	-0.0208**	-0.0173**
	(0.00820)	(0.00828)
4 years after reform implementation	-0.0260**	-0.0210**
	(0.00863)	(0.00875)
5 years after reform implementation	-0.0259**	-0.0194**
	(0.00906)	(0.00924)
6 or more years after reform implementation	-0.0200**	-0.0192*
	(0.00996)	(0.00925)
After national reform implementation in 2006		0.00908***
		(0.00234)
Time trend	0.00350***	0.00189***
	(0.000570)	(0.000662)
Constant	-6.984***	-3.759***
	(1,141)	(1.325)
Observations	8486	8,486
R-squared	0.313	0.316
Number of municipalities	107	107

NOTE: Standard errors clustered at municipality level. Before municipality implementation indicates all years prior to implementation. After national implementation means the years after 2006. Municipality fixed effects. I am controlling for the use of assistants, student-teacher ratio, enrolment, school type and the proportion of male teachers that varies over time.

First, the time trend coefficient is positive, which indicates positive growth in special education for non-reformed municipalities. As we can see from figure 2, non-reformed municipalities seem to have a stronger growth than reformed municipalities. The findings from the JLS analyses confirm this point, as shown by the point estimates for the dummy variables, which are all negative and significant. As discussed in detail above, the point estimate for D^k is interpreted as the change in special education placements for reformed municipalities k years after the reform, relative to the change for non-reformed municipalities. The point estimate is thus the change in the special education gap between reformed and non-reformed municipalities. An increasing negative coefficient indicates that this gap is increasing, i.e., the growth in special education placements is lower in reformed municipalities than in non-reformed municipalities. For the years prior to implementation, this difference is between 0,875 percentage points and 1,23 percentage points. For the years after reform implementation, this difference is between 1,39 percentage points and 2,6 percentage points. From the third year onwards it is above 2 percentage points for all years. The reason for the negative effect before implementation most likely indicates that the governing system in the reformed municipalities likely differed from that of the non-reformed municipalities before reform implementation, which is consistent with the representation in figure 2. One of the advantages of this strategy is that I can monitor these initial effects and identify the effects that are caused by the change in governing systems following the reform implementation year. In the years after the reform, the growth in the special education level is significantly lower in reformed municipalities than in the non-reformed. The first years after reform implementation, the coefficient does not change significantly. However, after 3 years, the coefficient is more than twice as large as before the implementation year. Thus, 3 years after the reform implementation year, the coefficient is approximately 1,4-1,6 percentage points

larger than it was in the years before implementation. The reason that the coefficient does not increase significantly in the first years after implementation is most likely because reform implementation takes time and the rest of the organisation need time to adapt the new systems. Additionally, it may be related to long-term effects being larger than short-term effects. The last dummy coefficient is smaller than many of the others. This result may occur because the municipalities that implemented reforms more than 6 years ago were not particularly influenced by the national reform and these reforms differ from the reforms implemented in the years following "Kunnskapsløftet". Additionally, the last reform dummy has few observations and may, therefore, be vulnerable to outliers.

As discussed above, there are several potential problems with my strategy. First, this strategy does not consider the effect of the national reform in 2006. There should be a dummy variable included that captures the effects of the national part of the reform. A dummy variable that is 1 for observations in the period after 2006 and 0 otherwise is included in Table 3, in the second column. It is significant and positive. However, including this variable does not affect the rest of the analyses that are worth mentioning. All coefficients are reduced somewhat in magnitude, but the difference before and after reform implementation is approximately the same. There are other issues that remain. The effects of the municipal reform might differ before and after national implementation. Thus, interaction terms should be included to capture this relationship. Furthermore, the analyses reported above are all school-level analyses. The reason for choosing school-level analyses is that the decision of providing special education to students is most often made by the school leaders. However, because the governing systems are decided at the municipality level, I include a model applying the JLS-strategy at the municipality level. These specifications have fewer observations (including municipality-fixed effects); therefore, one should not expect significant point estimates for

each year in the model. However, the magnitude of the effects should be comparable. Finally, the JLS strategy implicitly assumes that the municipalities that have not implemented the reform yet will not do so in the near future. This issue may create bias in my estimates because some of the municipalities that are in my reference groups should be included in the dummy categories for the years before implementation. Excluding the municipalities with the lowest probability of implementing reforms should address this problem.

Table 4 takes into account two of the remaining problems. First, I have included a time dummy, which captures the effect of the national reform, just as in table 3. To analyse if the effects of the municipal reform differ before and after the national implementation, I add together several reform dummies. I generate a dummy variable “before” for the years before implementation. I generate a dummy variable “reform” for the reform implementation year and the first two years following reform. Finally, I generate a dummy variable “after” for all the rest of the years after the municipal reform implementation. The results are given in table 4.

Table 4: Limited JLS analyses with interactions with the national reform. The level of special education as dependent variable

	Level of special education	
Before municipality implementation	-0.00145	-0.00145
	(0.00293)	(0.00295)
Municipality reform implementation years	-0.00416	-0.00257
	(0.00302)	(0.00292)
After municipality implementation years	-0.00962**	-0.00830**
	(0.00448)	(0.00391)
After national reform implementation	0.00939***	0.00975***
	(0.00193)	(0.00265)
Interaction between municipal reform implementation years and the national implementation period		-0.00207
		(0.00389)
Interaction between the years after municipal implementation and the national implementation period		0.000355
		(0.00253)
Time trend	0.00163***	0.00160***
	(0.000585)	(0.000578)
Constant	-3.236***	-3.178***
	(1.172)	(1.158)
Observations	8,486	8,486
R-squared	0.314	0.314
Number of municipalities	107	107

NOTE: Standard errors clustered at municipality level. Before municipality implementation indicates all years prior to implementation. Municipality implementation years indicates the implementation year and the two following years. After municipality implementation indicates the third years after implementation and onwards. After national implementation means the years after 2006. Municipality fixed effects. I am controlling for the use of assistants, student-teacher ratio, enrolment, school type and the proportion of male teachers that varies over time.

The first column in table 4 is almost similar to the JLS analyses in table 3. We have a significant negative estimate for the dummy variable capturing the period after reform implementation, and the point estimate for the dummy variable for the time period after national implementation is positive. The interaction terms are not significant, which indicates

that there are no systematic differences in the effects of the municipal reform implementation before and after national implementation. As discussed above, this might be as expected because elements of the new governing system were introduced prior to national implementation of the reform, which saw both the abolition of class size regulations and the introduction of the national tests. The rest of the coefficients in table 4 confirm my initial JLS analyses. However, table 4 is less flexible with fewer variables.

Table 5 addresses the two other issues presented above. First, in column 1, I perform an analysis at the municipality level because governmental systems are decided at the municipality level; the municipality-level analyses function as a robustness check. Including municipality-fixed effects, we only exploit variations for one municipality over several years. However, although the magnitude of the effects is approximately the same, only a few coefficients are significant. These findings confirm the results from my initial JLS analyses – the coefficients are less precise because of fewer observations.

My JLS analyses implicitly assume that a municipality that has not yet implemented the accountability reform will not do so in the coming years. If they were to implement the reform, they should instead be included in the reform dummies before implementation and not in the comparison group, as shown in table 3. Above, I discussed that the reason we find an effect before implementation is that the municipalities' governing systems also differed before implementation. The municipalities that implemented a full accountability system – with leader contracts and systematic evaluations – were most likely closer to an accountability system before implementation than those municipalities that did not implement the reform at all. Therefore, I suggest that those municipalities that have either leader contracts or systematic evaluations but do not have a full accountability system have a higher probability

of implementing a full accountability system in the coming years than municipalities that have neither leader contracts nor systematic evaluations. Because of this pattern, I exclude the municipalities with the highest probability of implementing reforms from the comparison group. The reference group then consists of municipalities with neither leader contracts nor systematic evaluations. Column 2 in table 5 addresses this problem. The coefficients are larger than reported above with approximately the same precision. The difference before and after implementation is also slightly larger. This finding should not come as a surprise. The comparison group has changed from being a group that included municipalities with a very low probability of reforming and municipalities with a higher probability of reforming to a group of municipalities that have only a low probability of reforming. The new comparison group is characterised as the less-reformed municipalities with neither leader contracts nor systematic evaluations. According to my empirical investigations, these municipalities will have a substantial increase in special education placements because of the national aspects of the reform, most likely because they do not have the right tools to handle the increased demand for special education. Therefore, the coefficients both before and after the municipal implementation, should be larger after excluding these observations. The results in table 5 are confirmative. However, the conclusion holds despite this difference.

Table 5: Municipality level analyses and robustness check. The level of special education as dependent variable.

	Level of special education	
	Municipality level analyses	**Excluded Municipalities with high probability of implementing reforms**
4 years or more before reform implementation	-0.00972*	-0.0120***
	(0.00550)	(0.00437)
3 years before reform implementation	-0.0100*	-0.0153***
	(0.00556)	(0.00566)
2 years before reform implementation	-0.00921	-0.0170**
	(0.00564)	(0.00677)
1 year before reform implementation	-0.0120**	-0.0196**
	(0.00575)	(0.00801)
Reform implementation year	-0.0122**	-0.0205**
	(0.00618)	(0.00933)
1 year after reform implementation	-0.0148**	-0.0240**
	(0.00710)	(0.0105)
2 years after reform implementation	-0.00725	-0.0255**
	(0.00844)	(0.0127)
3 years after reform implementation	-0.0131	-0.0337**
	(0.0143)	(0.0130)
4 years after reform implementation	-0.0273	-0.0402***
	(0.0175)	(0.0142)
5 years after reform implementation	-0.0243	-0.0415***
	(0.0175)	(0.0153)
6 or more years after reform implementation	-0.0246*	-0.0387**
	(0.0142)	(0.0175)
Trend	0.00327***	0.00493***
	(0.000434)	(0.00155)
Constant	-6.503***	-9.843***
	(0.872)	(2.310)
Observations	849	5225
Number of municipalities	107	55
R-squared	0.276	0.358

NOTE: Standard errors clustered at municipality level. Municipality fixed effects. Column 2 excludes municipalities with a partly implemented accountability system. I am controlling for the use of assistants, student-teacher ratio, enrolment, school type and the proportion of male teachers that varies over time.

Conclusion

I have discussed Norwegian institutional settings in the context of the three-tier principal-agent framework with asymmetric information. Norway has a federal governing system. It is the local councils who decide the number of schools, their locations, their governing systems and their budgets. The national government constrains the municipalities through different. Until the years immediately before the "Kunnskapsløftet" reform, the national governance system was characterised by many strict national rules and regulations that significantly influenced the operations of the educational sector. With the national reform, many of these rules and regulations were abolished and national tests were simultaneously introduced.

Accountability systems at the national level might increase the demand for special education because the teacher and/or parents have stronger incentives to acquire more resources for the class. The principal-agent framework presented in this paper predicts that municipal implementation of accountability reforms will decrease information asymmetry between the school owner and teachers. This decreased asymmetry will, in turn, give the school owner more power to handle the increased demand for special education. However, there is evidence that municipalities have implemented the reform to varying degrees and at different points in time. This finding gives rises to the generalised difference-in-difference model, introduced by Jacobsen, Lalonde and Sullivan (1993).

By using the JLS strategy, I find that the increase in special education placements is lower in reform-implementing municipalities than in non-implementing municipalities. The difference between the level of special education between reformed and non-reformed municipalities is larger after the reform implementation; the difference is approximately 1,6 percentage points. I have made these analyses both at the school level and at the municipality level and find

comparable coefficients. However, I do not find any systematic differences in the effects of the municipal reform from before and after the national implementation.

Paper III

CONDITIONAL PEER EFFECTS: THE ROLE OF SCHOOL AUTHORITIES

Jon Marius Vaag Iversen

Centre for economic research and Department of Economics, NTNU

N-7491 Trondheim, Norway

Abstract

The current empirical literature on peer group effects highlights that credible peer effects cannot be estimated unless parental sorting is taken into consideration. The present paper, which uses data from the Norwegian elementary schools, takes this existing empirical literature as its point of departure. The contribution of this paper is that in addition to parental sorting, it highlights the role of school authorities when estimating gender peer effects. Specifically, I focus on how special education assignments are strategically used by the authorities and the impact of this strategy on the estimated peer effects. I analyze the role of special education and investigate how principals use special education resources to affect the educational setting, the learning environment and the estimated peer effects.

Keywords: educational economics, resource allocation

JEL classification: I20, I28

E-mail: jon.iversen@svt.ntnu.no

Tel: +47 73 59 19 72

Introduction

The notion that peer effects are important to educational outcomes has great intuitive appeal. Peer effects have thus long been a concern to social scientists, educators and parents, and they are presently one of the most active areas within the economics of education. Peer groups are characterized in several ways in the literature, for example regarding the proportion of high achieving students or immigrants in a group, the educational level among parents at the school, or gender composition of a group. In short, there is currently a large empirical literature on peer effects, but little is known about the causal impacts of peers owing to methodological issues as discussed in Manski (1993).

Exogenous variations within schools over a period of time have been used as an empirical strategy to estimate gender peer effects since Hoxby (2000). Recent examples are Lavy and Schlosser (2010), Black, Deveraux and Salvanes (2010) and Bifulco, Fletcher and Ross (2011). This literature emphasizes solutions to difficult endogeneity problems regarding parental sorting but does not discuss how the school authorities may act to affect the peer composition. However, a reasonable assumption is that the school actors observe these externalities and have opportunities and possible controlling mechanisms to affect the educational setting and the peer composition. Peer group manipulation by school actors is thus likely to be a general problem when estimating peer effects. Lavy and Schlosser (2010), who investigate gender peer effects, argue that school controlling mechanisms or compensating resources are not problematic in their analyses when including a school-specific linear time trend. In addition, for their study, it is noted that Israeli students in need of special education are given instruction in separate classes at separate schools. However, this issue is of particular interest for studies of countries with institutions such as those in Norway.

The Norwegian elementary school is characterized by full inclusion policies and the generous use of special education resources. In this paper, I examine one observable form of the school actors' actions: the special education assignment.

The use of special education in the Norwegian elementary schools is characterized by boys receiving 70 percent of these resources, and there is some evidence that their overrepresentation in special education is due to behavioral problems. These boys might cause negative peer effects and have a negative impact on the learning environment (Nordahl and Hausstätter (2009)). Moreover, in the last decade, the proportion of special education students has doubled, and it is still increasing – implying that gender peer effects might unfold differently today than, for ten years ago. Johnsen, Nergaard and Flaatten (2006) have shown that the number of segregated groups receiving special education has grown as much as 300 % the last 15 years. Intuitively, segregated groups will be the most effective means for reducing the negative peer effects from these students, and this historical development is important for my analyses. The purpose of the present paper is to use data from Norwegian elementary schools to investigate whether, or to what extent, the estimates of gender peer effects are conditional on the level of special education resources.

Major methodological problems arise when including principals and school owners in the study. A reasonable assumption is that the schools themselves decide whether to use special education resources to solve behavioral problems. If a student does not benefit from ordinary instruction, he or she has the right to receive special education treatment. In theory, there should therefore not be any constraints on special education. However, special education is a very expensive way of teaching, and at some point, there will be a financial constraint. All teaching is taken from the same school budget. In practice, there is a distinct possibility that

low-quality schools are characterized by being less able to manage boys and using large amounts of special education resources. This compensation will attenuate some of the negative externalities at these schools. Nevertheless, I address the following questions: are the gender peer effects conditional upon actions from the school authorities, and can we identify important tools that school authorities can use to affect the educational setting? If the school actors use Special education to attenuate the negative externalities from disruptive boys, we must take this behavior into account when estimating the peer effects. However, the unobservable differences between schools can make it difficult to analyze the relationship between peer effects and the actions of school authorities.

I begin by analyzing gender peer effects while taking account of parental sorting. In these analyses, I only include the proportion of students receiving special education as a control variable. I proceed by focusing on the relationship between gender peer effects and the proportion of students receiving special education in more detail. I demonstrate that the gender peer effect is conditional on the level of special education. When the level of special education increases, the gender peer effect on student achievement is reduced significantly. To analyze to what extent the gender peer effects are conditional upon the actions from school authorities, I continue to perform comparisons between schools based on their level of special education. Ideally, I should compare the same students in different settings. However, that strategy is not possible with the existing data. Step by step, I try to make this comparison as credible as possible.

My best attempt, and the most credible strategy, is to exploit variations over time; hence, I concentrate my analyses on schools that have special education students at the 5^{th} grade for one or two years, but no special education for at least one year. Within this setting, I compare

various points of time within the same schools. I discuss all of these strategies and their implications in Section 5.

Using Ordinary least squares (OLS), I find significant negative effects for the proportion of boys in the 5th grade. Adding fixed school effects yields no significant effects for all students at the 5 % level over a period of three years. My results indicate that the gender peer effects are conditional upon the level of special education. In schools with no special education, the gender peer effects are relatively strong and significant. When school authorities intervene by providing special education to one or more students, my peer estimates are reduced significantly. These findings indicate that schools have succeeded in reducing negative externalities by using special education. I identify special education as an important tool for reducing externalities and positively affecting the educational setting.

The structure of the paper is as follows: Section 2 discusses some theoretical contributions to this issue; Section 3 introduces the data, characteristics of the Norwegian school system and my empirical strategy. Section 4 presents my results for all peer effects, special education and the robustness checks, while Section 5 concludes the paper.

2. Important literature and theoretical considerations

Many of the early empirical investigations of peer effects were not much guided by theory. In more recent literature, Lazear (2001) treats classroom education as a public good and introduces externalities as an important factor in educational production. When one student disrupts a class, learning is reduced for all of the other students. I present the main arguments of Lazear (2001) and some implications of this theory for my empirical analyses.

The key variable in Lazear's model is *p*, which is the probability that a student will behave well, such that *1-p* is the probability that a student will interrupt. When all students in a class of size *n* are behaving in the same way, disruption occurs $1-p^n$ of the time. More disruption means less time for instruction. This definition, of course, is a very stylized characterization of the classroom.

Lazear's theory can be used to understand the actions of school authorities. Students in the same class will most likely have a different *p* to represent their behavior. For principals and school owners, the student's *p* may be observable, and they will act to reduce the consequences from having students with a low *p*. Removing, or taking care of the students with the lowest *p* by providing special education increases the instruction time in the classroom, p^n. This increase may come from two sources: 1) Providing special education may reduce *n*. Removing or taking care of a student may reduce the group size, thereby allowing more time for instruction. 2) Providing special education to a student with a low *p* will allow more time for instruction, either by increasing this particular student's *p* or by taking the student with a low *p* out of the class, making the expression of p^n larger.

Guided by Lazear's theory, a small number of recent analyses focus on the negative externalities deriving from subgroups of classmates. One important example is Lavy and Schlosser (2010), who use Israeli data to find that an increase in the proportion of girls in the class improves both the boys' and the girls' academic achievements; moreover, they provide evidence that the mechanism appears to be that a higher proportion of females in a class leads to a better classroom and learning environment. Students who have more female peers report a lower level of classroom violence and disruption and better relationships with other students and with teachers. In Israel, however, learning disabled students are taken out of the school

and placed in special education schools. The authors do not provide information on the selection of these students, or the criteria for removing them from ordinary instructions. Figlio (2003) and Fletcher (2010) are other examples of studies that find negative effects from attending a class with disruptive students. Figlio (2003) finds that disruptive classmates apparently reduce overall mathematics achievement and increase the likelihood that other classmates will become disruptive. Fletcher (2010) finds that students with classmates that have serious emotional problems, score significantly lower than other students. One contribution using Norwegian data is Bonesrønning (2006), who highlights the negative externalities related to classmates from dissolved families. All of these studies are silent about the use of special education.

Another gender peer study using Norwegian data is Black, Devereux and Salvanes (2010). These authors' focus is on the future outcome from gender peer effects, e.g., IQ scores at age 18, teenage childbearing, the post-compulsory schooling educational track, adult labor market status and earnings. They find evidence for gender peer effects, which differ between men and women.

I am unaware of existing empirical studies that highlight the potential importance of school actor behavior on peer group effects. Empirical investigations on the effects of special education resources are also limited, mostly owing to methodological issues. Even so, some studies focus on the effects of having classmates who receive special education, guided by the results in Hanushek, Kain and Rivkin (1998). These authors find that an individual with a larger proportion of classmates who receive special education will have higher achievement gains than other individuals. A study by Friesen and Krauth (2008) suggests that there are negative spillovers from classmates with special needs on academic performance.

The current paper deviates from the existing gender peer group literature in several important ways. First, and most important, I emphasize the role of special education in estimating peer effects, because the relationship between classroom gender composition and student achievement is based on the level of special education. Through these exercises, I indirectly identify a potentially important tool, through which politicians, principals and administrators are able to attenuate negative externalities in the classroom. Second, I study the peer effects among young Norwegian students. The Norwegian school system is characterized by low mobility, which ensures that we have a good approximation of the students' peers throughout their school career up to the time of testing; i.e., the treatment is determined by good precision. Third, the combination of enrollments determined by location of residence and the small share of private schools (these do not depend on the neighborhood system) ensures that we are able to minimize the problems of students' endogenous mobility according to classroom gender composition.

In the following section, I present the data, the empirical strategy and my analyses of the gender peer effects and how these effects depend upon the level of special education.

3. Data, Empirical Strategy and institutional settings

Norwegian children enter the school system at the age of six, and all Norwegian children have the right to 13 years of education. Compulsory school can be discontinued at age 16, but in fact, approximately 95 % of the students proceed into upper secondary school for three years. The students are in primary school for their first seven years, in lower secondary school for the next three years and then in upper secondary school for the last three years. The primary and lower secondary schools are owned and fully financed by the municipalities, based on local taxes and national transfers.

Fewer than 2 % of Norwegian primary schools are private. Enrollment in a public school is determined by the location of the students' residence. When the student is six years old, he/she enters the neighborhood school together with other children from the same area. One implication of the neighborhood system is that the parents who want to change schools because of, for instance, a "bad draw in the gender composition lottery" experience high costs. Classes segregated by gender are not allowed in Norway, and the school organization is characterized by home classes, implying that the students spend almost all of their classes with the same peers. Early/late starting students and grade retention are extremely rare. Because of the neighborhood system and the school organization, we have a good approximation of the gender composition for their entire treatment period ($1^{st} – 4^{th}$ grade).

Every student in Norway has the right to special education if he or she does not benefit from ordinary instruction. The decision to provide special education to one or more students is made by the principal or school owner, although they need an expert evaluation from the educational psychological service, which is most often located in the municipality's administration (Utdanningsdirektoratet 2009). After this evaluation, the school owner or the school principal will make a decision about the amount of special education resources required by the student. In most cases, the principal makes this decision (Strøm et al., 2009). In any event, the parents should be involved throughout the entire process. In addition, a decision must be made about how the special education will be provided. The Norwegian system is characterized by full inclusion, meaning that all instruction should be carried out in the same classroom. There are, however, examples of students receiving special education alone or in small groups.

The principal has the opportunity to take special education students out of the testing pool for national tests. Nevertheless, there are reasons to believe that some of these students are still are tested. I would like to separate the pecial education students from the ordinary students in my analyses. The special education students are not identified in the data, implying that I cannot separate these students with as much precision as I would like. I address this cencern in the robustness checks.

3.1 The data

Starting in 2007, all of the 5th and 8th grades students in Norway are tested in the early autumn in math, reading in Norwegian and reading in English. I use the data from the tests taken during the 2007-2009 period, together with data from the schools (from GSI), families and municipalities. GSI is a school administrative system that collects information from the school principals for all primary schools and lower secondary schools in Norway. GSI covers data on teaching hours, number of students, special education, assistants, administration and so forth. Statistics Norway has supplemented these data with information about the students, their parents and the municipalities that they live in.

The data set has several nice features, and includes information about all Norwegian primary schools and all students entering 5th grade at all schools. These data, together with a long list of student and family characteristics for all students and a long list of characteristics for every single school, allow me to include many important controls in the regression models. I have chosen to focus on 5th graders, although I also have data on 8th graders. The reason for dropping the 8th graders is that most of these students enter in a new school when they begin their lower secondary schooling. For the students who change schools, I have no information about their peer and school history. I only have data for their peers for approximately a month

prior to testing, and I do not have any information about the school that they attended for the first seven years of their school career. Therefore there will be major issues with omitted variable bias if I try to estimate peer effects for these 8^{th} graders.

The outcome variable is constructed by summarizing the individual scores from national tests in all three subjects. The sum score is then standardized to a mean of zero and a standard deviation of one. The two explanatory variables of interest are the proportion of boys in the grade and the proportion of students receiving special education. Approximately half of the population is boys, but the between -schools and between-cohorts variation in the gender composition is still relatively high. Variations across schools are a little larger than variations within schools. Nevertheless, a within-school standard deviation of 0.08, and a total standard deviation of 0.12 indicate sufficient variation in both dimensions. While implementing this methodology, I use the proportion of boys measured at the grade level, not the classroom level, because I do not have information about the classes. However, this compromise may be preferable: classroom peer composition might be endogenous because parents and school authorities may have some discretion in placing students in different classes within a grade.

The level of special education in the school is measured by the proportion of all students in the 5^{th} grade that received special education during the year prior to testing. In other words, I use data for the 5^{th} graders from the year that they were in the 4^{th} grade. By using the level of special education the prior year, I overcome problems with schools that have increased their level of special education after the students are tested. Approximately five percent of fifth graders receive special education. The percentage is a little higher in small schools than in larger schools. Table 1 provides important descriptive statistics on the dependent variable and

the two primary explanatory variables. The table divides the samples into groups by school size and whether the school has special education students.

Table 1: Descriptive statistics for the dependent variable and the main explanatory variables of interest. Average values

Sample Variables	All students	Small schools	Large schools	Schools with no special education	Schools with special education
Proportion of boys	0.51	0.51	0.51	0.5	0.51
Proportion of students receiving special education	0.051	0.058	0.044	0	0.067
Standardized values of national tests	0	-0.05	0.09	0	0.01

Table A1 from the Appendix provides essential descriptive statistics for the explanatory variables mentioned above as well as for other important control variables. Using data for all three years, we have a total of 149,514 students. Not surprisingly, approximately half of the population is girls. The father's education is given as NUS-codes, with a value of 4 corresponding to three years of upper secondary education, a value of 5 corresponding to four years of upper secondary education and a value of 6-8 corresponding to higher education. On average, both fathers and mothers receive education corresponding to 3-4 years of upper secondary school, and approximately 74 % of the students live together with both of their biological parents. The average student enters a grade with an enrollment of roughly 40 students.

3.2 Empirical Strategy

The identification of the peer effect is not straightforward, with a primary concern being the effect of unobserved correlated factors. Much of the literature focuses on correlation from parental sorting or self-selection across schools due to differences in gender composition. In the first part of the empirical analyses, I address with these challenges by closely following the approaches that are established in the literature. In the second part of the empirical analyses, I evaluate whether the estimated peer effects are affected by the school actors' behavior. The main strategy here is to make comparisons that are as credible as possible between schools at different levels of special education.

Addressing parental sorting

To account for the challenges related to the self-selection of schools by parents, I rely on within-school variations and exploit variations in the proportion of boys over time within the same school. Using this approach, I investigate whether cohort-to-cohort changes in the achievement level within the same grade and school are systematically associated with cohort-to-cohort changes in the proportion of male students. In following this procedure, I apply the assumption that the students face the same school environment except for differences in gender composition. I will return to address the credibility of this assumtion in the second part of this section.

My initial empirical investigation thus follows the strategy introduced by Hoxby (2000). Using repeated cross-sectional data, I estimate the Equation (1) below:

(1) $A_{ist} = \beta_s + \beta_1 P_{st} + \beta_2 F_{ist} + \beta_3 X_{st} + u_{ist}$

where *i* denotes the individuals, s denotes the schools and *t* denotes time. A_{ist} is an achievement measure of student *i* in school *s* and year *t.* β_s is a school effect and P_{st} is the proportion of male students in school *s* at time *t.* F_{ist} includes characteristics by the individual student and his family. This vector contains information about gender, immigrant status, parental education, parental income, family type and birth order. X_{st} is a vector of the school characteristics in school *s* at time *t,* and it contains information about the grade size, the school educational level, the assistant resources, the proportion of teachers without approved education, the proportion of male teachers and the level of special education; these characteristics vary over time within the same school. The coefficient of interest is β_1, which captures the effect on achievement of having more male peers in the same grade.

Addressing teacher and school principal behavior

The second part of the analyses is motivated by the observation that the gender composition is correlated with the level of special education resources. I estimate an equation using the proportion of students receiving special education as the dependent variable and the proportion of boys as the explanatory variable of main interest. After controlling for factors at the individual, family, and school levels, the regression model in Table 2 indicates that the proportion of boys has a positive and significant effect on the proportion of students receiving special education resources – this relationship holds even when the school fixed effects are included in the estimated equation. This relationship could simply reflect that boys are overrepresented in the group of special education students. To address this concern I also estimate the same model using the proportion of girls receiving special education as the dependent variable, and my models demonstrate a significant relationship here as well. These

findings indicate that we must take into account the use of special education when estimating the gender peer effects from equation (1).

Table 2: The relationship between the proportion of boys and the proportion of students receiving special education resources.

	Proportion of students receiving special education resources				
	OLS	OLS	OLS	OLS	FE
Proportion of boys	0.038***	0.039***	0.039***	0.037***	0.033***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.007
Proportion of students from intact families			-0.018***	-0.021***	-0.019**
			(0.002)	(0.002)	(0.009)
Educational level among fathers at the grade		-0.014***	-0.0088***	-0.0044***	-0.0019
		(0.0002)	(0.0003)	(0.0003)	(0.003)
Income level among fathers at the grade			-2.57e-08***	-1.80e-08***	-2.57e-09
			(1.13e-09)	(1.09e-09)	(9.03e-09)
School fixed effects	No	No	No	No	Yes
Constant	0.027***	0.086***	0.090***	0.069***	0.057***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.013)
Observations	145,547	145,547	145,547	145,547	145,547
R-squared	0.013	0.039	0.045	0.085	0.020
Number of schools					2,428

Note: The set of covariates is specified in the text, under the data presentation. Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants. Robust standard errors. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Equation (1) provides estimates that rely on the assumption that there are no variations in cohort level factors from one year to the next except for the student body composition. This assumption is not necessarily fulfilled if the principals or the school owners provide some students with extra resources such as special education resources. Table 2 indicates that this concern might be of importance. A simple, but regrettably naïve, solution to this problem is to include an interaction between the proportion of boys and the proportion of students with special education, as included in equation (2).

$$A_{ist} = \beta_s + \beta_1 P_{st} + \beta_2 F_{ist} + \beta_3 X_{st} + \beta_4 SE_{st} + \beta_5 SE_{st} P_{st} + u_{ist} \quad (2)$$

There are two new elements in this equation. First, I include a term for the time varying special education level, taken out of the X-vector. I also include an interaction term between the peer variable and the special education term. If the point estimates $\beta_3 - \beta_5$ are significant, it might indicate that the estimated peer effects depend on the use of special education.

However, this approach does not address the potential endogeneity of special education resources. Special education resources might be correlated with the residual in equation (2) for at least two reasons: we do not observe the student behaviors that qualify for special education, and we do not observe the teacher characteristics that influence the likelihood that a student misbehaves in class.

My best strategy is to exploit the within-school variations over time and concentrate on the schools that have had special education students in the 5^{th} grade for one or two years but no special education for at least one year. More precisely, I concentrate my analyses on the schools that have variation in the level of special education that is equal to 0 one year and that has a positive value the next year. I compare various points of time for the same schools. The

simple idea underlying this approach is that in the years with no special education, the estimated gender peer effects are not affected by the school actors' actions, and in years with special education, the school actors have affected the peer estimates, and we expect to find smaller effects. However, the variation in special education across years might be non-random. It appears likely that the learning environment is better in a year when no special resources are allocated than in the years when some resources are allocated to the actual cohort. However, if this situation is realistic, the identical environments would create even larger differences in gender peer estimates. I will discuss this atrategy and others more fully in Section 5.

Even if a within-school approach would overcome the potential problems regarding parental sorting, I conduct several robustness checks and diagnostic analyses, to support my analyses and to investigate whether there are other empirical concerns in estimating gender peer effects. Most of these checks and analyses follow those of the main critics of peer effects estimation in the literature.

4. Results

I begin by presenting the results for gender peer effects, taking account of parental sorting. Thereafter, I introduce the challenges associated with the use of special education resources. All of the analyses are based on equations 1 and 2.

4.1 Estimating gender peer effects while accounting for parental sorting

Table 3 reports the initial results from the OLS and the fixed effects models based on equation (1). In the first two columns, I only include the controls for individual characteristics and family background, while in the last three columns, I include school fixed effects. In the last

two columns, I also include controls for school factors that vary over time, whereas in the last column, I include the proportion of students who receive special education.

In the first two columns without the fixed effects, the point estimates for the proportion of boys are negative and statistically significant at the conventional levels. When including the school fixed effects in Column 3, only exploiting the within-school variations for a short period of three years, the peer coefficient is negative and significant at the 10% level. The result is very robust for including school factors that may vary over time. Based on the largest estimate in column 1, an increase in the proportion of boys from 40% to 60% will affect achievement by roughly 2% of a standard deviation. This effect is smaller than that found by Lavy and Schlosser (2010), who report an approximate increase of 4-5% of a standard deviation from a 20 percentage point change in the proportion of boys.

Table 3: Gender peer effects for the entire population of students in the cohorts 2007-2009.

	Standardized values of national tests 2007-2009				
	OLS	OLS	FE	FE	FE
Proportion of boys	-0.103***	-0.104***	-0.0802*	-0.0828*	-0.0689
	(0.0257)	(0.0243)	(0.0453)	(0.0451)	(0.0451)
Proportion of students receiving special education					-0.407***
					(0.0888)
Control for individual characteristics	x	x	X	x	x
Control for family background		x	X	x	x
School fixed effects			X	x	x
School controls				x	x
Control for special education					x
Constant	0.0948***	-0.979***	-0.898***	-0.902***	-0.884***
	(0.0141)	(0.0168)	(0.0262)	(0.0873)	(0.0871)
Observations	145,547	145,410	145,410	145,401	145,401
R-**squared**	0.004	0.126	0.096	0.096	0.097
Number of schools			2,428	2,424	2,424

Note: Robust standard errors clustered at the school level are reported in parentheses. Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants. Columns 1 and 2 present OLS results. Columns 3-5 include school fixed effects * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

The theory by Lazear (2001) and the empirical findings in Lavy and Schlosser (2010) imply that the mechanisms by which the peer effects operate are driven by negative externalities from disruptive students. In the appendix table A5, I present analyses where I break down

subgroups of boys by parental education. This separation is motivated by the findings in Haraldsvik and Bonesrønning (2012) that students with less educated parents are associated with greater negative externalities. The result from these analyzes indicate there are more negative externalities related to boys with less educated parents. I find no negative effects from the share of boys with highly educated parents.

4.2 Estimating gender peer effects while taking account of the school actor's actions

As shown in column 5 in table 3, when controlling for special education in the regression model the coefficient for the proportion of boys in the grade becomes smaller and loses its significance. The proportion of boys and the proportion of students receiving special education resources are positively correlated, so this result is not surprising (see Table 2). I have already established that the use of special education should be taken into account when estimating peer effects. Table 3 does not fully address the problem that the gender peer effect depends on the level of special education. The actions of the school authorities will most likely reduce the negative externalities related to boys in the classroom, and the problem might go away or be removed from the group of boys to the group of students receiving special education. I now present some approaches to estimating the role of school authorities, and the impact of the estimated peer effects.

Table 4 provides some introductory analyses for the interaction between the gender peer effects and the use of special education. Column 1 reports the results from estimating equation (2) presented in the empirical strategy. In the first column, I report the point estimates for three terms: 1) The proportion of boys; 2) the proportion of students deemed eligible for special education, and 3) an interaction between these two terms. The coefficients reported in column 1 indicate that there is a significant negative effect on achievement in

cohorts with a low proportion of boys. Yet, as the level of special education increases, the coefficient for the proportion of boys becomes less negative. Note also that the estimate for the proportion of students with special education is significantly negative. These results are open to interpretation. The results might indicate that special education is used by the school principal to reduce the negative externalities associated with boys and that the principal is moving the problem away from the group of boys into the group of special education students. Most important, these results suggest that the gender peer effects are conditional upon actions from school authorities,which we must account for in the empirical strategy. The purpose of the following analyses is to evaluate the consequences of these actions on the estimates of the gender peer effects.

Column 2 in table 4 reports the results from the non-special education case, that is, when all schools that have special education are excluded from the sample that year. We see that in a situation with no interventions, the point estimate for the proportion of boys and achievement is larger than it is for the case with all schools included, and it is significant at the 10 % level. One worry about the specification in Column 2 is that the included schools differ too much from the group of schools with a positive level of special education, and that these differences may cause the differences in peer estimates. For instance, schools without special education may use other controlling mechanisms to improve the educational environment. Alternatively, the schools with no special education may have teachers who are better able to handle student misbehavior. Table A2 in the Appendix provides some evidence that this assertion may reflect the actual situation. The difference in observable characteristics between the two groups of schools is not large, but the students in schools who use special education are slightly more likely to have educated parents, and higher family income. In addition, these schools are also smaller, likely because small schools are less likely to offer special education. As the number

of students in a grade increases, the probability of having a student receiving special education increases as well. Because of this concern, I present another specification where I keep the schools without special education but narrow the focus of the sample even further by only keeping the schools with a total size of between 50 to 150 students. The reason for keeping only these schools is that they probably share some of the same characteristics, as demonstrated in Table A2 in the Appendix. The results which are reported in column 3 from table 4 show an even larger estimate for the proportion of boys, that is more comparable to the results found in Lavy and Schlosser (2010). A 20 percentage point change in the proportion of boys is related to a change in achievement of 5.5% of a standard deviation. The coefficient is significant at the 5% level.

Table 4: Comparing schools with some of the same characteristics, and dividing them between schools with no special education and schools with special education. School fixed effects.

Standardized values of national tests 2007-2009				
Sample:	**All students**	**Schools without special education students**	**Schools *without* special education students. Enrollment 50-150 students**	**Schools *with* special education students. Enrollment 50-150 students**
Proportion of boys	-0.110**	-0.131*	-0.283**	-0.109
	(0.0531)	(0.0783)	(0.139)	(0.135)
Interaction between the proportion of boys and the proportion of students receiving special education	0.721*			
	(0.433)			
Proportion of students receiving special education	-0.793***	0	0	-0.386
	(0.255)	(0)	(0)	(0.272)
Constant	-0.862***	-0.577***	-0.903***	-0.932***
	(0.0888)	(0.161)	(0.229)	(0.230)
Observations	145,401	33,920	9,185	11,869
R-squared	0.097	0.095	0.081	0.096
Number of school	2,424	1,566	596	599

Note: Robust standard errors clustered at the school level are reported in the parentheses. Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants. All columns reports results on school fixed effects models.
* significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

There is still a concern, however, that the schools with different levels of special education also differ in other observable or unobservable ways, even when I compare small schools to other small schools, and that these differences are the reason for the different peer estimates. Ideally, one should compare the schools with themselves, and I have the possibility to make this type of comparison when using time variations. Some schools inconsistently offer special education during 2007-2009. The solution that I present here is to focus the analyses on all of the schools that had some special education in one or two of the years, but no special education in at least one year. More clearly, I concentrate my analyses on the schools that have a variation in the level of special education from 0 in one year to a positive level the next year. The schools that have no special education students in all three years do not contribute to these analyses, nor do the schools with a positive level of special education in all three years. We then compare the peer coefficients for the same schools during the years that the do not use special education resources to the peer coefficients when they have a positive level of special education. In this setting, I compare the schools with themselves at various points in time. All of the schools contribute to both parts of the sample, making the comparison credible. The teachers are the same, the principal is probably the same and both the school and the municipality policies about special education are constant.

Table 5: Comparing schools with themselves. Schools included have special education students for one or two years, but no special education students for at least one year. OLS.

Standardized values of national tests 2007-2009		
	With special education students	**No special education**
Proportion of boys	-0.0507	-0.151***
	(0.0484)	(0.0472)
Level of special education	-0.271***	0
	(0.0991)	(0)
Constant	-1.366***	-1.211***
	(0.0508)	(0.0526)
Observations	35,775	28,110
R-squared	0.138	0.127

Note: Robust standard errors clustered at the school level are reported in the parentheses. Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants. All columns reports results on school fixed effects models. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Because we have the same schools in both groups, the comparisons of background variables in Table A2 show identical samples. The mean values are almost identical in terms of family background, school resources, and the number of municipality inhabitants, with the only difference being the size of the schools. The school enrollment and the grade enrollment appear to be larger in the years with special education and I control for this in my regressions.

When the schools do not have any special education students, we estimate significantly negative gender peer effects (see table 5). When the level of special education is positive, we have no significant peer effects. This finding indicates that special education has an effect on the educational setting, and that it works in a way that attenuates the gender peer effect. The effects are also stronger than in the models that include the entire population. In this case, a 20 percentage point increase in the proportion of boys decreases student achievement by a little more than 3 % of a standard deviation.

Some concerns remain, that need to be discussed. Why do the schools use special education some years and not others? One important reason could be the differences in teacher quality from one year to another within a school. One could think that the quality of the teacher is higher in the years where the schools have no special education versus the years where the schools do provide special education. However, if this is the situation, identical teacher quality would make the differences in peer estimates even larger. The relatively large initial difference in my best approach, however, suggests that this problem is negligible.

The variation in the student composition from one year to the next may also cause bias in my estimates, for similar reasons. In groups with more capable and students that are less disruptive, there will be no need for special education. If this is the case, identical students would result in even larger differences in peer estimates, as in the case with teacher quality. Thus, the unobserved differences in teacher quality and in student composition work in the same direction and could create a downward bias in my estimates. However, the comparison presented here would reduce these problems to a minimum, and thus, these problems would not cause significant bias in the estimates.

Analyzing a sample of schools in this manner raises questions about external validity. This sample is not randomly drawn or representative drawn out of the population. Moreover, the sample is taken out of the population based on observable characteristics, more precisely based on the level of special education. One concern is that these schools differ from the rest of the population, and that they are, in that sense, not generalizable. A comparison across columns in table A2, suggests that this concern is insignificant.

School owners and principals have a strong management tool and an important policy instrument that can improve the educational setting. It is crucial to take this tool into account

in situations when estimating the peer effects. However, there are still problematic students in the school system. Some of these students are taken out of the classroom, while some remain and are provided with special education services inside the classroom. The coefficient for the proportion of special education is difficult to interpret. Even so, this issue is appropriate for further research. My results show that the peer effects are conditional upon actions from the school authorities.

4.3 Robustness checks

There are some concerns regarding my empirical specification. First, there are other aspects of the relationship between special education and gender peer effects that should be taken into account. Most important, the relationship between these two variables could result from non-linearity in the gender peer effects. Because the boys dominate the special education group, the special education term could capture the non-linearity, and we would, in fact, be estimating a more flexible gender peer effect. Table A3 presents a model where the special education term is excluded. Instead, I have included a squared term of the gender peer term. This model does not support the hypothesis that the peer effects are non-linear.

Another concern is that the special education students are included in the dependent variable. In fact, most of these students are excluded. Principals have the opportunity to exclude all students with special education from the testing pool. But, in case the opposite has occurred, I present models in table A3 that correct for special education students by excluding the students with the highest probability of receiving special education. My results are robust for these specifications.

I continue by running a falsification check. I use the proportion of boys in the 4th grade instead of the proportion of boys in the 5th grade as explanatory variable. The proportion of boys in the 4th grade should not have any effect on achievement for the 5th graders. These results are reported in Table A3 and are satisfying. The proportion of boys in the 4th grade has no effect on the achievement of the 5th graders.

As suggested by Lavy and Schlosser (2010), I run a balancing test. Even if fluctuations in the proportion of boys within a school resemble a random process, these variations could in fact be correlated with additional cohort to cohort changes that may affect student outcomes. To assess this possibility, I check whether the changes in the proportion of boys within a school is associated with changes in the student background characteristics, such as parental income, education, family structure and so on. Table A3 provides the results of these tests. We see, as expected, that the proportion of boys is not related to any of the student characteristics. However, one could still be concerned that the students might respond to unpredicted shocks in student composition. The neighborhood system, the lack of school choice in Norwegian primary schools, and the limited scope of private schools significantly diminishes this concern.

Inspired by Altonji, Elder and Tabor (2005), I also test the robustness of our peer estimates by progressively adding different types of observable student covariates to my regression models. As illustrated in table 3, adding individual or family covariates does not significantly change the coefficient. Following the intuition in Altonji et al., this result indicates that the potential bias from unobservables is small. However, when including controls for time-varying school variables and school fixed effects, the coefficient becomes smaller and more imprecise. This issue will be a central part of my paper.

I also construct a peer proxy that corrects for the individual students' sex. In my earlier analyses, the individual student's sex contributes to the gender composition. This contribution could create a small bias in my estimates as well. By correcting for the individual students' sex, my conclusions hold. The coefficient and the significance hardly changes. This evidence is also provided in table A3.

Although not reported in the tables, I find that the peer effects are not a phenomenon related to small schools or small municipalities, by dividing the total sample by the school size or by the number of inhabitants in the municipality. My results are not affected by the inclusion of school specific time trends, and the school level analyses provide results with identical conclusions.

5. Conclusions

Much of the peer group literature assumes that the peer effects are unconditional upon the school actors' actions. This paper presents evidence that this assumption could be restrictive. This paper discusses the role of special education in estimating gender peer effects, and I identify classroom externalities that derive from differences in gender composition. At the outset, I follow the empirical strategy introduced by Hoxby (2000). For this short period of time, I find no significant gender peer effects for the entire population of early elementary school at conventional significance levels. In estimating these models for a period of three years, I neglect the attempts of the school actors to address the potential negative externalities. However, I present reasons to believe that the peer estimates are conditional upon the use of special education and that increasing the level of special education attenuates the negative peer effects from the share of boys.

I find that in the schools with a positive level of special education, the gender peer coefficients are small and insignificant, while in the schools with no special education, the gender peer effects are negative and significantly different from zero. In one interpretation, these findings indicate that the school actors apply special education resources in ways that reduce the negative externalities produced by the boys. Nevertheless, the question is whether the externalities are eliminated or whether they have simply moved to a different group of students, i.e., the special education students. For the reasons discussed earlier, it is not possible to interpret the negative coefficient for the proportion of students receiving special education. To estimate the impact of peers' special education on achievement, we need a credible instrumental strategy, which is an issue for further research.

The contributions of these analyses are twofold. First, they highlight that gender peer effects are conditional upon the school actors' behavior. Second, the analyses provide evidence that special education has an effect on other students in the classroom. These contributions have strong policy implications. By providing these groups with extra resources through special education for students with behavioral problems, the principal might be able to reduce the negative externalities. One major limitation of these analyses is that I am unable to establish the size of the gender peer effects net of the special education resources.

Appendix I

Table A1: Descriptive statistics

VARIABLE NAME	MEAN	STANDARD DEVATION
Proportion of boys at the school level	0.511	0.12
Proportion of first-generation immigrants	0.01	0.1
Proportion of second-generation immigrants	0.037	0.19
Father's education	4.4	1.63
Mother's education	4.53	1.62
Proportion of students in intact families	0.74	0.44
Grade size	39.9	21.4
Proportion of students provided special education	0.051	0.053

Table A2: Descriptive statistics for the groups of schools with and without special education

	All students		Enrollment between 50 and 150 students at the school		Schools dropping in and out of the group with special education	
	No special education	**With special education**	**No special education**	**With special education**	**No special education**	**With special education**
Mother's education	4.47	4.54	4.25	4.24	4.5	4.53
Father's education	4.31	4.42	4	4	4.34	4.38
Intact families	75.3%	73.8%	76,7%	74.6%	75	74.4%
Mother's earnings	263 000	271 000	239 000	236 000	267 000	269 000
Father's earnings	485 000	494 000	428 000	425 000	492 000	492 000
First generation immigrants	0.8%	1.1%	0.4%	0.4%	0.9%	1%
Second generation immigrants	2.5%	4.1%	0.8%	0.8%	2.8%	3%
Proportion of boys	50%	52%	50.1%	52.1%	50%	52.3%
Assistants	14.8	16.1	15.9	19	15.1	15.9
Uncertified teachers	7.2%	6.6%	7.1%	6.1%	7.1%	6.6%
Male teachers	25.1%	24.4%	24.6%	24.9%	24.9%	24.9%
Enrollment	140	219	90.1	98.7	154	172
Grade enrollment	17.2	28.3	11	13.4	18.8	22.5
Municipality inhabitants	11 784	11 684	13 366	13 608	12 305	12 585

Table A3: Robustness checks

Standardized values of national tests for fifth grade students 2007-2009						
					Sample reduced by students with high probability of receiving special education	
	All students	All students	All students	All students	All students	No special education
Proportion of boys	-0.139	-0.148			-0.0779	-0.211***
	(0.199)	(0.199)			(0.0494)	(0.0561)
Interaction		0.710				
		(0.436)				
Squared term	0.0688	0.0383				
	(0.189)	(0.190)				
Falsification test.			0.0261			
			(0.0420)			
Peer measure corrected for individual student's sex				-0.104**		
				(0.0487)		
Constant	-0.867***	-0.853***				
	(0.0993)	(0.100)				
Observations	145,401	145,401	145,401	145,401	103,724	19,936
R-squared	0.097	0.097	0,097	0,097	0.096	0.109
Number of schools.	2,424	2,424	2,424	2,424		

Note: Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants..
Robust standard errors. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Table A4: Balancing test: The relationship between gender composition and pre-determined student attributes

	School fixed effects models					
Dependent variable	**Father's education**	**Mother's education**	**Dissolved families**	**Mother's earnings**	**Father's earnings**	**First generation immigrants**
Proportion of boys	-0.0140	0.0170	-0.00263	-5,297	10,805	-0.00253
	(0.0432)	(0.0442)	(0.0133)	(5,359)	(10,954)	(0.00274)
Constant	2.266***	2.621***	0.357***	65,648***	142,871***	0.0245***
	(0.0505)	(0.0490)	(0.0141)	(5,932)	(15,137)	(0.00319)
Observations	145,401	145,401	145,401	145,401	145,401	145,401
R-squared	0.189	0.243	0.050	0.120	0.048	0.016
Number of schools	2,424	2,424	2,424	2,424	2,424	2,424

Note: Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants..
Robust standard errors. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Appendix II

Breaking the group of boys into subgroups by parental education

The theory by Lazear (2001) and the empirical findings in Lavy and Schlosser (2010) imply that the mechanisms through which the peer effects operate relate to negative externalities from disruptive students. Here, I present analyses where I break out the subgroup of boys by parental education. This separation is motivated by the findings in Haraldsvik and Bonesrønning (2012), that the students with less educated parents are associated with negative externalities. The results from these analyses indicate there are more negative externalities related to the boys with less educated parents. I find that there are no negative effects from the share of boys with highly educated parents.

Thus, in this case, the peer variable from Equation 1 will, in turn, be the proportions of boys with less educated and highly educated parents. I have divided the groups by the average parental education, such that in the group of less educated parents, none of the parents have any education beyond upper secondary school. The results are reported in Table 4.

In column 1, the coefficient for the boys with less educated parents is twice as large as the initial peer coefficient. To evaluate the size of the effect, we have to take into account that, changes in the proportion of boys with less educated parents are smaller than the changes in the overall proportion of boys. Thus, a change of 20 percentage points in the proportion of boys with less educated parents changes the average achievement by 2.5% of a standard deviation. As can be seen from column 2, the effect of a change in the proportion of boys with highly educated parents is minimal and insignificant. It appears clear that I have managed to trace one group of boys who affect other students more than others. This finding confirms my own hypothesis, the theoretical implications from Lazear (2001) and the empirical findings in Lavy and Schlosser (2010) that the mechanisms through which the peer effects operate relate to negative externalities from disruptive boys. Obviously, school authorities will try to affect the educational setting to minimize the negative externalities coming from the boys. Therefore, the gender peer effect will be conditional upon these actions, and more emphasis should be placed on this relationship than is seen inthe existing literature. My gender peer

effect coefficients in table 3 will be downward biased because of this pattern and I proceed by presenting some evidence on the magnitude and the sign of this bias.

Table A5: Peer effects of being enrolled together with boys with a higher probability of being disruptive. Different definition of peer groups in the models.

Specification of peer: Proportion of students who are:	**Boys**	**Boys from dissolved families**	**Boys with less educated parents**	**Boys with more highly educated parents**
	Dependent variable: Standardized values of national tests			
Peer group	-0.0689	-0.0725	-0.124**	-0.00614
	(0.0451)	(0.0702)	(0.0607)	(0.0857)
Proportion of students receiving special education	-0.407***	-0.413***	-0.405***	-0.418***
	(0.0888)	(0.0888)	(0.0886)	(0.0887)
Interaction between the educational level at school and the proportion of boys				
Constant	-0.884***	-0.902***	-0.839***	-0.920***
	(0.0871)	(0.0845)	(0.0935)	(0.0852)
Observations	145,401	145,401	145,401	145,401
R-squared	0.097	0.097	0.097	0.097
Number of schools	2,424	2,424	2,424	2,424

Note: Controls include student gender, immigration status, parental education, parental income, family structure, number of brothers and sisters, birth order, grade enrollment, school type, proportion of male teachers, level of special education and number of assistants.. Robust standard errors. * significant at 10 % level. ** significant at 5 % level. *** significant at 1 % level.

Paper IV

ARE NON-ELIGIBLE STUDENTS AFFECTED BY SPECIAL EDUCATION

Jon Marius Vaag Iversen**, Hans Bonesrønning*, and Ivar Pettersen**

*Norwegian University of Science and Technology, and Norwegian Business School BI

** Center for Economic Research at NTNU

Abstract.

We investigate whether the academic performance of non-eligible students – in an institutional setting of full inclusion - are affected by special education resources. Using data from the Norwegian elementary school, evidence is provided that the academic performance of non-eligible students is negatively correlated with the proportion of students eligible to special education, presumably because misbehaving students are overrepresented within this subgroup. The hypothesis to be investigated is that more hours of special education per eligible student might improve the academic performance of non-eligible students, presumably by dampening the negative externalities. We take advantage of a large, across-the-board increase in the proportion of eligible students, and combine fixed effects with an IV-approach to identify the causal effects of special education on the academic performance of non-eligible students. We find that non-eligible students are positively affected by an increase in *the number of hours in special education per eligible student*.

JEL-codes: I20, I28, H52

Key words: student achievement, special education, externalities

INTRODUCTION

Special education is specially designed instruction that aims at improving the performance of students, who because of a disability, do not benefit from ordinary classroom teaching. The purpose of the present paper is to investigate whether such resources have non-intentional effects by benefitting non-eligible students. This issue, which at first thought might seem strange, is motivated by a recent, large increase in the proportion of students that are deemed eligible to special education in the Norwegian elementary school. Many of the "new and non-core" special education students are deemed eligible due to misbehavior in class. There is quite a heated discussion whether these students should be offered special education within class or out of class. See Bateman (1996) and Crockett and Kauffman (1999) for early overviews of the US discussions of this issue. Norwegian national guidelines state that all students benefit from inclusion, and thus that special education should be offered within ordinary classrooms. This view is challenged by people who argue that non-eligible students are negatively affected by inclusion, and that - at least some - misbehaving students might benefit from being treated by specialists, in separation from well-behaving students. Neither party can back their point of view by much empirical evidence. The purpose of the present paper is to contribute to this discussion by providing one piece of empirical evidence: conditional upon full inclusion policies, do special education resources improve the performance of non-eligible students?

The most likely mechanism linking the performance of non-eligible students to the presence of eligible students is identified by Lazear (2001), who conceptualize education production as a public good with congestion. That is, effective teaching time is determined by the probability that the students misbehave in class. Many of the core special education categories

of students are most likely *not* associated with high probabilities of misbehavior. Students that have visual or hearing impairments fall into this category. When special resources are allocated to these types of students, non-eligible students might be affected by a more generous student-to-teacher ratio. Existing evidence indicates that any positive effects associated with a more generous teacher to student ratio are likely to be small. When special education resources are allocated to students that are eligible to special education due to behavioral problems, the special resources might dampen or eliminate misbehavior/negative externalities, and thereby increase effective teaching time and improve teacher-student relationships. In such cases, special education resources might potentially have significant positive effects on the performance of non-eligible students.

There is quite a lot of evidence from around the world that negative externalities are present in classrooms. In the US, 85 percent of teachers, and 73 percent of parents say that "school experience of most students suffers at the expense of a few chronic offenders" (Public Agenda, 2004). The present study uses Norwegian data. This country ranks top in the PISA-studies when it comes to student misbehavior. For instance, 40 percent of the 15 years old students that participate in the PISA-studies report that all or most of the lessons are characterized by disruptions and noise. However, these investigations do not link disruption to particular subgroups of students. Turning to empirical analyses, Figlio (2007), Bonesrønning (2008), Carrell and Hoekstra (2010), Fletcher (2010) and Lavy and Schlosser (2010) provide evidence that negative externalities are related to "boys named Sue", students from dissolved families, students that have experienced domestic violence or are mentally unstable, or simply are boys. Many of the students under scrutiny in these analyses might be exposed to treatment by special education resources, but none of the studies cited above raise the issue whether the

negative externalities are affected by such treatment. Carrell and Hoekstra (2010) come closest when they worry that their results will be biased towards zero if the level of domestic violence in a school-grade-year is correlated with common shocks. The allocation of special education resources might be an example of such a shock.

Empirical studies showing how non-eligible students are affected by the presence of eligible students, or by special education resources, are rare. We are aware of just two earlier contributions. Friesen, Hickey and Krauth (2010) find that attending schools with a higher percentage of students with learning disabilities or behavioral disorders has a small negative impact on the reading and numeracy test scores of non-disabled students. Hanushek, Kain and Rivkin (2002) find that the achievement growth for non-eligible students is positively related to the proportion of special education students. They investigate a number of potential mechanisms, but alternative specifications only reinforce their central result that disabled students do not harm the academic achievement of their peers.

Estimating the effects of special education on the performance of non-eligible students entails several econometric problems. Perhaps the most obvious is that the amount of special education reflects unobserved characteristics of students, teachers, school principals and school owners (the municipalities). We deal with these challenges by exploiting a strong increase in the proportion of eligible students following the introduction of a national education reform with accountability elements in 2006. By using adjacent cohorts of students, we introduce school-by-grade level fixed effects to get rid of time-invariant between-school variation in special education that reflects unobserved teacher and school principal quality, and municipality fixed effects to get rid of time-invariant unobserved school owner characteristics. It is recognized that a substantial part of the remaining variation in special

resources - across adjacent cohorts within schools - reflects unobserved student and teacher quality. To get rid of this potentially endogenous variation, the amount of special resources is instrumented. Thus, the econometric approach combines fixed effects with instrument variables.

To increase the transparency of the analyses, the rigorous part of the paper is preceded by some exploratory analyses. The purpose is to highlight why, and under what conditions, special resources might affect the performance of non-eligible students. This part provides evidence that those students who are exposed to classmates that are deemed eligible to special education report relatively more noise and disorder, and that the performance of individual students is negatively associated with the proportion of eligible students in the grade. Importantly, the exploratory part also provides indicative evidence that the negative association between the proportion of eligible students and the achievements of non-eligible students is weaker when more special resources are allocated per eligible student, and when the eligible students are offered their own lessons in separation from non-eligible classmates. The latter findings might potentially indicate that the negative externalities decrease or disappear when the misbehaving students are segregated out of ordinary classrooms. Consistent with these findings, the instrumental variable, fixed effects analyses provide significant positive estimates for the amount of special resources per eligible student on the performance of non-eligible students.

The rest of the paper is organized as follows. The next section provides some descriptive statistics, and notably, a description of the treatment variables. Thereafter we present the exploratory analyses and the more rigorous analyses. The last section offers some concluding remarks.

DATA, DESCRIPTIVE STATISTICS AND MEASURES OF TREATMENT

In this section, we first briefly describe our data, and go on by describing the construction of the three key variables that are used to characterize treatment, which are the number of hours in special education per student, the number of hours in special education per eligible student, and the share of eligible students.

Data

The Norwegian Government introduced nationwide tests in mathematics, reading in Norwegian and English language for 5th graders in 2007 as part of an accountability reform. We take advantage of these tests and use administrative records for three consecutive cohorts of 5th grade students (2007-2009) in the Norwegian elementary school. Statistics Norway has linked the national test results to individual (gender, ethnicity, birth order) and family characteristics (mother's and father's education, mother's and father's income, family size, family structure) for the entire population of 5th graders. Detailed information about school enrollment and school inputs is provided by the national Elementary School Information System (GSI). Our information about special education resources comes from this source. The national tests are taken early in the fall of the 5th grade and we use information about school inputs for the preceding school year, that is, when the students were in the 4th grade. In addition, we exploit data from a yearly survey provided to all students from the 5th grade and onwards to assess the classroom climate. No descriptive statistics for the individual and family background characteristics are presented in the paper, but such information is available upon request.

Our outcome measure is generated from the national test results. The 5th grade students sit mandatory tests in mathematics, reading in Norwegian, and reading and writing in English in the start of the fall semester. The tests have different scales. We have standardized the tests, added the results, and then standardized once more. The outcome measure thus has a mean of 0 and standard deviation equal to 1. It is an issue whether these practices of aggregating across different subjects conceal important between-subject differences. Investigations show that reporting separate results for each subject add little additional insights, perhaps because the Norwegian elementary school system practices home classes.

All non-eligible students sit the tests (if not absent on the days the test are taken), while eligible students can apply for exemption. Clearly, the school actors' incentives are to encourage special education students to apply for exemption because this is an easy way to improve on average student achievement. We know that approximately 10 percent of the students do not participate in the tests, which is substantially above the proportion of students that receive special education. In 2008 the participation rate varied from 78.6 to 95.1 percent among the 19 counties in Norway. The county with the lowest participation rate had a proportion of eligible students equal to 7.9 percent this year, while the county with the highest participation rate had a proportion of eligible students equal to 4.2 percent. In all counties the participation rate is lower than the proportion of non-eligible students. Unfortunately, the eligible students are not identified in the data, and we do not know the number of eligible students that have applied for exemption. Although it is likely that most of the special education students are among the 10 percent of students that do not sit the tests, we cannot say this for sure. In most analyses we include all students that have participated in all three tests.

The potential biases these practices introduce to our analyses are evaluated by first regressing the share of eligible students against student body characteristics, and thereafter using the information from this analysis to exclude from the analyses the student subgroups that most likely receive special education resources.

Special resources

The Norwegian elementary school is embedded in a federal system, where multi-purpose municipalities (about 430) run the public elementary and lower secondary schools (a total of about 2900) subject to national laws and regulations. The municipalities are financed by local taxes – tax rates set by the national government - and national grants. Special education is handled within this system as follows. The right to special education is regulated by national law. The law says that students that do not benefit from the ordinary teaching are entitled to special education. Entitlement is determined by experts hired by the municipalities. Eligible students are assigned to one of the following categories: visual or hearing impairment, communication problems, brain damages, learning disabilities, concentration problems, or misbehavior (related to ADHD, other diagnoses or no specific diagnoses). Having received a diagnosis, eligible students are assigned a total number of hours in special education per year. This decision is made within the municipality. In a survey to the municipalities in 2009, about half of the municipalities answered that this decision was taken at the municipal level, while the other half answered that this decision was decentralized to the schools. In the former case, the municipal officers allocate the total educational budget in the municipalities to the local school, determining the allocation of resources between special education and alternative uses for each of them. In the latter case, the schools face a within-year fixed budget, and have to allocate their resources between alternative uses. Our econometric specifications, to be presented below, reflect this variation in decision-making authority.

The organization of special education is guided by the principle of full inclusion. Thus, most students that are deemed eligible are taught in ordinary classes by adding a special education teacher or an assistant. Alternatively, the special education students are taught in smaller groups of eligible students for a limited number of hours, or sometimes tutored (alone) for a limited number of hours. Only students with the most serious kinds of retardation are taught in special schools. In 2001 2.1 percent of the students in the capital of Oslo and 0.4 percent of the students in the rest of the country were enrolled in special schools.

From 2006 and onwards the Elementary School Information System ("GSI") has reported the amount of special education resources by grades in schools. Taking account of the existing institutions, we derive three measures from these statistics; the proportion of students that are deemed eligible, the number of hours in special education per eligible student by grades in schools, and the product of these two measures which is the number of hours in special education per student in the grade. We have also considered a fourth measure; which is the fraction of eligible students that are taught in groups with other eligible students, that is, separated from non-eligible students.

Descriptive statistics are reported in Table 1. In 2006 5.6 percent of the students in the 4^{th} grade were deemed eligible students. By 2008, this proportion had increased to 7.4 percent of the 4^{th} grade students, implying that the proportion of eligible students has increased by more than 30 percent in the three-year period. An important feature is that boys are much more likely to be deemed eligible than girls. Roughly, there are 3 eligible boys per each eligible girl.

Table 1 Descriptive statistics. Measures of treatment

	2006	2007	2008
Proportion of eligible students	5.65	6.42	7.42
#hours of special education per eligible student	156	152	146

For 4th graders the number of hours in special education per eligible student per year declined from 156 hours in 2006, to 152 hours in 2007 and 146 hours in 2008. This decline is not sufficient for the special education budget to stay fixed throughout the period. Additional resources are provided – in principle these resources might come from an expanding educational budget or from reallocation of resources, that is, special education is financed by cutting back on ordinary teaching, building maintenance or other inputs. It seems that the major part of increase in special education hours per student is financed by expanding the total municipality budgets to education. Elsewhere we have shown that the municipalities' incomes have increased in the period, and that the number of hours in special education per student is income elastic, that is, the number of hours increases when municipalities' incomes increase. The raw data show that the number of students per (ordinary) teacher has not decreased in the actual period: in 2006 and 2007 there were on average 9.9 students per teacher, in 2008 the number was 10.0. As an additional exercise, we have investigated whether ordinary teacher man-years are substituted for special education. This is done by regressing special education hours per student against ordinary teaching hours per student (both variables measured at the municipality level) while controlling for a number of time-varying municipality characteristics and municipality fixed effects. For the period we are investigating here, the point estimate for ordinary teaching hours per student is -0.15 and

highly insignificant, indicating that there is no statistical significant input substitution between ordinary teacher man-years and special education resources. Thus, it basically seems like the decision makers face an increasing budget, but are forced to make some trade-off between the proportion of eligible students and the number of special education hours per eligible student.

While the expansion of special education has not affected the number of teachers per student, it is paralleled by an increase in the proportion of uncertified teachers, from 2.4 percent in 2006 to 4.1 percent in 2008.

As mentioned above, we have also considered a direct measure of segregation. The percentage of boys that receive special education in groups with other eligible students – that is, segregated from their non-eligible classmates - has increased slightly, from below 4 percent in 2007 to 5.5 percent in 2009. National experts on special education have told us that these statistics are less reliable than the other statistics on special education. Their argument is that the school principals' incentives to report correct numbers to the national data base differ across items: the organization of special education is a controversial issue, so reporting high numbers of students that are treated outside their home classes might lead to potential sanctions by national government bodies, while the numbers of students and hours to special education are part of the school budget implying that the numbers reported to national data base have to be identical to the numbers that appear in the budget.

Here we therefore report analyses that use the number of hours in special education per eligible students to characterize one dimension of treatment. Regrettably, the number of hours per eligible students is a somewhat noisy and imprecise measure of segregation because some schools use these resources to separate out the special education students completely, while others do not. Thus, some schools might achieve a lot more segregation if they lump hours of

special education together and provide special education for a group of students instead of providing it on a one-to-one basis within the same classroom as ordinary students. We have investigated whether the less reliable measure of segregated special education provides different results, but this is not the case. These results have the same flavor to them, but are less precise than those presented below.

We also would like to know how special resources are allocated across different categories of eligible students. Notably, we would like to know the proportion of misbehaving students within this subgroup. The Elementary School Information System provides no inform about this, so we have been searching for information elsewhere. First, the Norwegian Institute for Public Health (NIPH) states that by 2005 1.1 percent of the Norwegian population less than 18 years has an ADHD diagnosis. There is an overrepresentation of children aged 12-15 years and four out of five with an ADHD diagnosis is a boy. Further, NIPH states that the proportion of the population with an ADHD-diagnosis is rapidly increasing, reflecting that about 3-5 percent of the young population actually carry this disease. In the elementary school, students with an ADHD-diagnosis are automatically classified as being eligible to special education. Second, in two surveys to more than 2000 elementary school teacher in 2006 and 2008 (Nordahl and Hausstätter (2009)) classified about 10 percent of the boys and 2 percent of the girls to the two categories "students with behavior problems, but not ADHD" and "students with ADHD". We have performed a regression analysis with the proportion of eligible students at the grade level as the dependent variable, and student body characteristics as independent variables. The hypothesis underlying this exercise is that core special education categories as visual and hearing impairment are randomly allocated across gender and family background characteristics, while misbehaviors are not. Finding that special education is associated with student and family characteristics are thus an indication that this

student subgroup contains misbehaving students. All the independent variables are aggregated to the grade level, and we have used data from three years. Table 2 reports the results.

Table 2 Determinants for the proportion of eligible students

VARIABLES	Proportion of eligible students-grade level	Proportion of eligible students-grade level	Proportion of eligible students-grade level	Proportion of eligible students-grade level	Proportion of eligible students-grade level
Proportion of boys	0.038***	0.039***	0.039***	0.037***	0.033***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.007
Proportion of students from intact families			-0.018***	-0.021***	-0.019**
			(0.002)	(0.002)	(0.009)
Father's education – garde average		-0.014***	-0.0088***	-0.0044***	-0.0019
		(0.0002)	(0.0003)	(0.0003)	(0.003)
Father's earnings – grade average			-2.57e-08***	-1.80e-08***	-2.57e-09
			(1.13e-09)	(1.09e-09)	(9.03e-09)
Constant	0.027***	0.086***	0.090***	0.069***	0.057***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.013)
Observations	145,547	145,547	145,547	145,547	145,547
R-squared	0.013	0.039	0.045	0.085	0.020
# schools					2,428

Note: *** p<0.01. All specifications include year dummies. The two specifications to the right also include the number of students in the grade.

Column 1 includes the proportion of boys as the only independent variable together with year dummies. Additional student body characteristics are included in columns 2-4, and school fixed effects are added in column 5. The point estimate for the proportion of boys is strongly significant throughout the table; indicating that an increasing proportion of boys in the grade causes more eligible students. Also, there are indications that the proportion of eligible students is systematically related to socioeconomic characteristics of the family. The point estimate for the proportion of students from intact families is significant throughout the table. This finding is consistent with the empirical literature on family dissolution, which provides evidence that children from dissolved families are more likely to misbehave compared to students from intact families (e.g. Ermisch and Francesconi (2001) and Painter and Levine (2000)). The estimates for father's education and earnings are significant in specifications that exploit all kinds of variation, but become insignificant when fixed school effects are included. Running the equations reported in Table 2 separately for the proportions of eligible boys and girls respectively, it is evident that only the proportion of eligible boys is significantly associated with the proportion of intact families (not reported in tables). These latter findings are consistent with Bertrand and Pan (2011) who report that boys' non-cognitive skills are negatively affected by growing up in a single parent family.

EXPLORATORY ANALYSES: SPECIAL EDUCATION AND NEGATIVE EXTERNALITIES

In this section we report results from exploratory analyses that will be helpful in making sense of the more rigorous, but less transparent, analyses that are presented later on. We start out by investigating whether eligible students carry negative externalities. Two exercises are provided towards this end. First, we investigate whether the occurrence of noise and disorder

in teaching situations, as reported by the students themselves, is related to the proportion of eligible students. Data from the national annual survey to students is used. The survey is voluntary for 5th and 6th grade students, and mandatory for 7th grade students, implying that the number of respondents is much higher in the 7th grade than in the 5th grade, and that analyses of the latter data are not clouded by selectivity problems. We therefore report results from using data from the 7th grade, but using the 5th grade surveys generates the same patterns. An indicator for the occurrence of noise and disorder is generated by combining the answers to two questions/statements to the students: "To what extent are you disturbed by misbehaving classmates during the work sessions?" and "I am often disturbed by other students when I am working at school." The noise and disorder indicator; which is increasing in the occurrence of disruption, is a grade level variable. It is regressed against the proportion of eligible students in the grade while controlling for socioeconomic characteristics of the student body. The point estimate for the proportion of eligible students is positive and significant at the 5 percent level, indicating that there is more noise and disorder in schools/grades where a large fraction of students are deemed eligible to special education. These results are reported in Appendix Table 1.

Second, we have investigated the relationship between student performance and the proportion of eligible students. The estimated equation has the standardized score for each student– aggregated over the three tests in Mathematics, reading in Norwegian and reading and writing in English – as its dependent variable and the proportion of eligible students in the grade as the independent variable of key interest. A number of controls at the individual student level - gender, ethnicity, birth order, parents' education and earnings, family size and family structure – are included together with a few school characteristics, such as school size, resource measures and other peer characteristics. This equation is estimated for each of the

three years for which we have data. As can be seen from Table 3, the number of students is somewhat below 50 000 students for each year. The cohort sizes are slightly above 60 000 students each year, implying that about 20 percent of the 5th grade population is excluded from the analyses. About 10 percent of the students do not sit the tests, and the rest 10 percent is excluded due to lack of information about individual and family background characteristics. As argued above, it is pretty safe to assume that very few special education students are represented in the regression analyses. The point estimates for the proportion of eligible students are negative and highly significant for all three years.

Table 3 The relationship between performance at national tests and the proportion of eligible students in the grade. 2007-2009.

VARIABLES	2007	2008	2009
Proportion of eligible students	-0.390***	-0.353***	-0.435***
	(0.089)	(0.086)	(0.081)
Observations	49,901	47,979	47,521
R-squared	0.133	0.133	0.128

Note: *** p<0.01. Control variables are gender, immigrant status, parents' education and earnings, family structure, family size and birth order, school size, peer group characteristics and measures of purchased inputs (teacher-student ratio, proportion of non-certified teachers).

Moreover, the negative effect associated with eligible students is of considerable size: an increase in the proportion of eligible students with 30 percent (equal to the increase in the proportion of students in the 2nd to 4th grade in the period 2006-2008) is associated with a performance decline of 0.12 standard deviations for non-eligible students. Our preferred interpretation of these estimates is that non-eligible students experience quite large negative effects from belonging to classrooms where many classmates are deemed eligible to special education. We are agnostic about why students are deemed eligible: This could be, for instance, because of an unfavorable student composition (the subgroup of eligible students

contains many misbehaving students), or because the teachers are of poor quality (which simultaneously leads to poor performance and many misbehaving students). At this stage, no attempts are made to sort out these explanations.

We have investigated whether these negative effects are dampened when more eligible students are treated in (more or less) isolation from non-eligible students, as indicated by the number of hours of special education per eligible student. For this purpose, the population of schools is separated into two categories; schools that use less than and more than the average number of hours of special education per eligible students, respectively. The education production function is estimated for both these subsamples using data for three year and a school fixed effects specification. The results are reported in Table 4.

Table 4 The relationship between student achievement and the proportion of eligible students in schools that use less and more than the average number of hours of special education per eligible student

	National tests			
VARIABLES	Less than the average number of hours per eligible student	More than the average number of hours per eligible student	Less than 100 hours per student per year	More than 200 hours per student per year
Proportion of eligible students	-0.603***	-0.0476	-0.590***	0.0517
	(0.138)	(0.172)	(0.186)	(0.235)
Observations	82,184	63,752	46,515	39,421
R-squared	0.101	0.093	0.106	0.092
#schools	1,753	1,816	1,191	1,413

Note: *** p<0.01. Control variables are gender, immigrant status, parents' education and earnings, family structure, family size and birth order, school size, peer group characteristics and measures of purchased inputs.

The point estimates for the proportion of eligible students are -0.603 and significant at 1 percent for schools that use less, and -0.048 and statistically insignificant for schools that use more, than the average hours of special education per eligible student. These results might *indicate* that schools, by excluding eligible students from the ordinary classrooms or from negative interactions with classmates, are able to reduce the potential negative externalities

that are associated with this student subgroup. We find similar results when the schools are partitioned into two groups based on the provision of segregated special education (not reported). These intuitive results are established without paying any attention to the inherent endogeneity problems. Nonetheless, this exploratory part motivates a hypothesis that non-eligible students might be positively affected by special education resources when the schools use these resources to increase the number of hours in special education per eligible student (which, for a fixed special education budget is equal to reducing the number of eligible students).

RIGOROUS ANALYSES: ARE NON-ELIGIBLE STUDENTS POSITIVELY AFFECTED BY SPECIAL EDUCATION RESOURCES?

The identification strategy

The exploratory analyses reported above indicate that special education resources are allocated in a compensatory way, that is, more special resources go to poor learning environments. To identify causal effects of special education resources on the performance of non-eligible students, we thus have to address the challenges related to two-way causality. We start out from the following equation:

$$A_{ijmt} = \alpha_j + \delta_m + \beta_1 SE_{jmt} + X_{ijmt}\beta_2 + Z_{jmt}\beta_3 + \sum \gamma_t D_t + \varepsilon_{ijmt} \quad (1)$$

where A_{ijmt} is achievement for the non-eligible student *i* in school *j* in municipality *m* in year *t*, α_j is a school-by-grade fixed effect, δ_m is a municipality fixed effect, SE_{jmt} is special education in school *j* in municipality *m* in year *t*. Inspired by the exploratory analysis we use three different measures of special education, the number of special education hours per student and its two components, which are the proportion of eligible students and the number of special

education hours per eligible student. X_{ijmt} is a vector of control variables at the individual student level, Z_{jmt} is a vector of time-varying school inputs other than special education and D_t are year dummies.

The municipality fixed effects eliminate the time-invariant between-municipality variation which is due to unobserved school owner characteristics and other time-invariant factors that are not included in the analyses. The school-by-grade fixed effects eliminate the between-school-by-grade-variation in the special education measures which is due to unobserved, time-invariant teacher and school principal quality. The remaining variation in special education resources across adjacent cohorts within schools reflects a combination of observed and unobserved student body characteristics, unobserved teacher characteristics, and increased generosity towards maladaptive students. Importantly, the proportion of students receiving special education has increased at all grade levels. As already mentioned, in the period 2006-2008 the proportion of eligible students in the 4th grade increased by 33.6 percent (from 5.55 percent points to 7.42 percent points), and in the grades 5-7 the proportion increased between 13.4 and 23.7 percent. We use this across-the-board increase in special education eligibility to facilitate identification. This is achieved by instrumenting the applied measure of special education for the 5th grade students (i.e. the special education they were exposed to in 4th grade) with the average special education used from the 5th to the 7th grade level in the same school in the same year.

All the three measures of special education at the 5th - 7th grades in the same school in the same year are highly correlated with the respective measures at the 4th grade. The validity of the instrument thus hinges on whether the exclusion restriction is fulfilled. There are no obvious mechanisms that mediate influences from the special resources in higher grades to the

performance of students in lower grades. At least, the application of school fixed effects, which remove the effects from time-invariant teacher quality, makes this less of a worry.

The results

All regression analyses presented in this section use the standardized score for individual students– aggregated over the three tests in mathematics, reading in Norwegian and reading and writing in English - as the dependent variable. In Table 5, treatment is measured by the number of hours in special education per student. As mentioned above, it is convenient to think of this measure as the product of the proportion of eligible students and the number of hours per eligible student.

Table 5 The casual effect of the number of hours of special education per student on the achievement of non-eligible students.

VARIABLES	OLS	FE_S	IV& FE_S	IV& FE_S& FE_M
#Hours of special education per student	-0.0043***	0.0004	0.0036**	0.0047**
	(-9.26)	(0.08)	(2.38)	(2,76)
Individual controls	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Observations	149622	149622	149451	148448
R^2	0.126	0.096	0.096	0.110

Note: Robust t-statistics in parentheses. *** p<0.01, ** p<0.05. Individual control variables are gender, immigrant status, mother's and father's education, mother's and father's earnings, family structure, family size, birth order.

Each cell in Table 5 shows the estimated coefficient on the special resources measure from a separate regression using data for three adjacent cohorts of students. Column 1 presents OLS estimates controlling for individual and family characteristics and year dummies. In column 2 school fixed effects are included. Column 3 is similar to column 2 except that the treatment

variable is instrumented. The specification reported in column 4 add municipality fixed effects interacted with time dummies to take care of potential influential time-invariant municipality factors that simultaneously affect student achievement and special education resources. We have interacted with year dummies to allow the unobserved municipality factors to exercise varying influences over the actual time period.

The point estimate for the number of hours in special education per student is negative and significant in column 1, close to zero and insignificant in column 2, and positive and significant in columns 3 and 4. The negative point estimate reported in column 1 indicates that there are negative externalities related to eligible students, or alternatively, that more special resources are used in classrooms with low quality teachers. These effects dominate any positive effects of special resources on student achievement. The change in the point estimate between columns 1 and 2 indicates that the negative estimate in column 1 incorporates between-school variation in unobserved student and teacher quality. Still, the column 2-estimate is most likely biased downwards due to within-school across-adjacent-cohorts variation in unobserved student and teacher characteristics. The IV-approach effectively takes away the within-school across-adjacent-cohort variation in special education resources that is due to unobserved student or teacher characteristics in those cohorts. Identification is thus based on the across-the board increase in special education in the period 2007-2009. The IV-estimate of 0.0036 reported in column 3 is nine times larger than the OLS-estimate reported in column 2; which we think is consistent with the existing empirical evidence that the within-school variation in teacher quality is likely to be quite substantial (see for instance Chetty, Friedman and Rockoff (2012)). That is, the sign of the OLS-bias is as expected because it seems likely that teachers of low quality simultaneously provide poor teaching and have disciplinary problems leading to an increasing number of diagnoses. The

point estimate reported in column 4 is based on the within-grades-in-school-within-municipality variation in special education where an instrument is used to remove the suspect variation in special education across adjacent cohorts, and is statistical significant and somewhat larger than the point estimate reported in column 3 – indicating that time-invariant unobserved municipality characteristics work to bias the point estimate downwards.

To evaluate the size of the effect reported in column 4, consider a school that has 40 students in the 5^{th} grade. We assume that two of the 40 students (5 percent) are eligible to special education, each with 150 hours per year. Then an additional 200 hours of special education hours are provided. At this stage, we do not care how these resources are used. It suffice to state that this equals an increase in special education of five hours per student; which, using the estimate from column 4, is transformed into (0.0047* 5=) 0.0235 standard deviations in academic performance for non-eligible students.

This seems like fairly small effects. However, there are reasons to believe that this analysis conceals the potential effects of special education resources on non-eligible students' performance. The exploratory analyses presented earlier indicate that the positive effect of special education, as reported in Table 5, depend on how the total number of special education hours is allocated among the students. This hypothesis is investigated by characterizing treatment by the number of hours per eligible student and the proportion of eligible students, respectively. These measures have to be included one at a time in the equation to be estimated due to strong partial correlation. Columns 1- 3 of Table 6 report the results from using the number of hours in special education per eligible student as the measure of treatment. The treatment variable is instrumented using the same approach as above, that is, we have used the number of hours in special education in the 5^{th}-7^{th} grades as our instrumental variable. School-

by-grade fixed effects, and municipality fixed effects interacted with year dummies, are included successively in columns 2 and 3.

Table 6 The casual effect of hours of special education per eligible student on the achievement of non-eligible students.

VARIABLES	FE_S	$IV\&FE_S$	$IV\&FE_S\&FE_M$	$IV\&FE_S$	$IV\&FE_S\&FE_M$
#hours per eligible student	0.0002	0.00019	0.00046*		
	(0.87)	(0.85)	(1.74)		
Proportion of eligible students				-1.43*	-1.75**
				(-1.66)	(-2.48)
Individual controls	Yes	Yes	Yes	Yes	Yes
School inputs	Yes	No	No	No	No
Year dummies	Yes	Yes	Yes	Yes	Yes
Observations	134363	134363	134363	146703	146703
R-squared	0.098	0.098	0,113	0.095	0.107
#Schools	2055	2055	1868	2390	2390

Note: Robust t-statistics in parentheses. *** p<0.01, ** p<0.05. Individual control variables are gender, immigrant status, mother's and father's education, mother's and father's earnings, family structure, family size, birth order.

The point estimate for the hours of special education per eligible student is positive in all the three specifications that are reported in the table, varying from 0.0002 and statistically insignificant in column 1 to 0.00046 and statistically significant in column 3. To evaluate the size of the latter effect we return to the school with 40 5^{th} graders, of which two initially are exposed to 150 hours each of special education per year. Assume once more that the school is provided with 200 more hours to use on special education, and is free to decide on the allocation of these resources. We consider the following alternatives. The school allocates equal amounts of the available resources (now a total of 500 hours) to the two existing special education students, that is, 250 hours per eligible student, or alternatively, the school allocates an equal number of hours to four students, that is, 125 hours per student. According to the point estimate in Table 6, column 3, the former alternative will, compared with the latter, lead

to (0.00046*125 =) 0.06 standard deviations better performance for non-eligible students. This example then illustrates that the effect of increasing special education resources per student might depend on how the additional resources are allocated across the student body.

In columns 4 and 5, treatment is characterized by the proportion of eligible students. We have instrumented this variable using the same approach as above. The point estimate is statistically significant and negative in both cases, indicating that the non-eligible students are worse off when the school classifies a large proportion of the students as eligible to special education.

As suggested above, the nearby explanation for the negative point estimate is that this is an indirect effect mediated through the budget constraint. The descriptive statistics presented earlier indicate that, even though the schools do not operate under fixed special education budgets, the two variables - the proportion of eligible students and the number of special education hours per eligible student - move in opposite directions. The negative point estimate for the proportion of eligible students in column 3 then echoes the positive point estimate for the hours of special education per eligible student in column 2. Using the point estimate in column 3, increasing the proportion of diagnoses from 5 (2 out of 40) to 10 percent (4 out of 40) while keeping the total special education budget constant – as in the example presented above – leads to a performance deterioration of (1.75*0.05=) 0.09 standard deviations. In absolute value, this effect is somewhat larger than the effect calculated above.

We could think of other mechanisms influencing the negative point estimate for the proportion of eligible students. For instance, diagnoses might have a stigmatizing effect, that is, misbehaving students might respond to a diagnosis by increasing their misbehavior. For instance, Morgan, Farkas and Hibel (2010) find that special education in some cases has a

negative effect on children's externalizing behaviors (which is a measure capturing the frequencies of arguing, fighting, showing anger, acting impulsive, and disturbing the classroom). These are not mutually exclusive explanations, but cannot be sorted out unless much more black box information is provided.

A robustness check

All the education production function specifications reported above might potentially include a small fraction of eligible students. These practices, which are dictated by the nature of the available data, potentially introduce a bias to our analyses. There are at least two problems to consider. First and much highlighted, some schools might use special education placements strategically to keep low performing students out of the testing pool. As long as these are time-invariant behaviors, the school fixed effects approach will take care of this. However, in the current data schools increase special education placements over time, implying that the estimates for the proportion of eligible students are biased upwards because there will be a "mechanical" positive relationship between the proportion of eligible students and student performance insofar that the students that are taken out of the testing pool and provided special education, perform below the population average. We have evaluated this bias by excluding from the analyses the student subgroups that are most likely to be deemed eligible to special education. From Table 2 it is evident that students with less educated parents and students from dissolved families are overrepresented among eligible students. We have therefore estimated all the relevant equations by first excluding students with parents that are educated at the lower secondary school or below, thereafter by excluding students from dissolved families, and finally, by excluding both students that have parents that are educated at the lower secondary level or below and students from dissolved families. Then we are left

with subgroups that are basically not affected by the increase in the proportion of eligible students, thus preventing that the estimated effects reflect changes in the composition of the treatment group. The point estimates change somewhat, but none of the results differ much from the results reported above. These results are not reported in the tables, but are available on request.

DISCUSSION AND CONCLUDING REMARKS

The main purpose of this paper has been to investigate whether special education resources affect the performance of non-eligible classmates. The data come from the Norwegian elementary school which is characterized by full inclusion, large proportions of eligible students and generous amounts of special resources. The investigations start out by providing evidence that there seems to be substantial negative externalities associated with the subgroup of students that are deemed eligible to special education, potentially reflecting that this subgroup contains a non-negligible fraction of misbehaving boys. Thus, the question addressed in the paper is actually whether special education resources work to dampen such negative externalities.

We have addressed the challenge related to the endogeneity of the special education resources by pooling data for three adjacent cohorts of 5^{th} graders and by combining school and municipality fixed effects with an instrumental variable approach, basically exploiting the rapid, across the board, increase in the proportion of students deemed eligible to special education in the period under scrutiny.

The evidence indicates that non-eligible students are positively affected by the number of hours of special education per eligible student, and negatively associated by an increase in the

proportion of eligible students. The favored interpretation of these results is that it matters quite a lot how the special education budget is allocated across the students. The following example illustrates. Making the assumption that all schools have the same size of 40 5^{th} graders and have 500 hours of special education hours to allocate, then, if all schools switch from allocating 125 hours to each of 4 students to allocating 250 hours to each of 2 students, the performance of the non-eligible students will improve by at least 0.06 standard deviations. This seems like a considerable effect following from a reallocation of resources within a fixed budget.

We end the paper by pointing to two issues for future research. First, it is not obvious that the school owners or school leaders can exercise much discretion with respect to the use of special education resources. In Norway, students who do not benefit from ordinary teaching, have a legal right to special education. There is a lot of anecdotal evidence saying that parents and teachers advocate eagerly for eligibility. The school owners or school leaders cannot overrule a decision of eligibility, but can adjust the number of hours per eligible student. The data used in this analysis reveal patterns of special education resources that are consistent with such behaviors. One topic for further research is how school leaders, and the local governments, seek to gain more control over the use and allocation of special education resources. A specific research question is whether local governments that have introduced accountability systems are more likely to exercise control over their special education resources.

The second research topic requires an opening-up of the classroom black box to access information about student and teacher behavior. The most likely mechanism underlying the findings reported in this paper is that non-eligible students benefit from the use of special

education resources because negative externalities are reduced: misbehaving students are taken out of the classroom or are taken care of by additional teachers within the classroom. In the present paper we have provided some scattered and indirect evidence that this might be the driving mechanism, but this topic would certainly gain from more thoroughly investigations.

Appendix

Appendix Table 1 The relationship between noise and disorder and the proportion of eligible students

	Noise and disorder	Noise and disorder	Noise and disorder
Proportion of eligible students	0.63***		0.63***
	(-2.73)		(-2.73)
# students	0.00036***	0.00031**	0.00036***
	(-2.69)	(-2.32)	(-2.69)
Proportion of boys	0.014	0.059	
	(-0.084)	(-0.35)	
Mothers' education – average	-0.069	-0.072	-0.070
	(1.42)	(1.45)	(1.43)
Fathers' education – average	0.064	0.059	0.064
	(-1.27)	(-1.17)	(-1.28)
Family structure – average	-0.38***	-0.38***	-0.38***
	(2.89)	(2.89)	(2.89)
Constant	6.28***	6.19***	6.27***
	(32.0)	(32.0)	(36.7)
Observations	2013	2013	2013
Adjusted R-squared	0.015	0.010	0.016

Note: *** p<0.01, ** p<0.05, * p<0.1 Robust t-statistics in parentheses

Appendix Table 2 The relationship between hours of special education per student and municipality size

VARIABLES	Hours of special education per student	Hours of special education per student
Method	OLS	FE
Population (ln)	-0.45***	-0.70***
	(-3.12)	(-2.92)
Revenues per inhabitant (ln)	6.06***	0.25*
	(7.09)	(1.69)
Share 0-5 yrs	-17.7	-2.18
	(-1.31)	(-1.42)
Share 6-15 yrs	-19.3**	-3.53***
	(-2.26)	(-2.81)
Share 67+	-18.5***	-2.42*
	(-3.10)	(-1.93)
Share 67-80 år	7.10***	-0.031
	(2.59)	(-0.094)
Scattered settlement	1.56***	0.12
	(3.04)	(0.55)
Observations	3346	3340
R^2_{adj}	0.210	-0.047
#municipalities		419

Note: *** p<0.01, ** p<0.05, * p<0.1 Robust t-statistics in parentheses

BIBLIOGRAPHY

Altonji, J. G., Elder T.E., and Tabor C.R. (2005) Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy,* 113: 151-184

Angrist, J. D. and V. Lavy, (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement, *The Quarterly Journal of Economics 114(2)*, 533-575.

Bateman, B. (1996). Better IEPs: How to develop legally correct and educationally useful programs. *Longmont*, CO: Sopris West.

Bertrand, M., and J. Pan (2011) The trouble with boys: social influences and the gender gap in disruptive behavior. *NBER Working paper #17541*

Bifulco, R., Fletcher J.M., and Ross, S.L. (2011) The effect of Classmate Characteristics on Individual Outcomes: Evidence from the Add Health. *American Economic Journal, forthcoming.*

Bitler, M.P., Gelbach J.H., and Hoynes H.W. (2006) What mean impacts miss: distributional effects of welfare reform experiments, *American Economic Review 96(4)*, 988-1012

Black, S.E., Devereux, P.J., and Salvanes K. (2010). Under Pressure? The Effects of Peers on Outcomes for Young Adults. NBER *Working Paper No. 16004*

Bland, M. (2000) An introduction to medical statistics, 3rd edition, *Oxford University Press*, Oxford

Bonesrønning, H., Borge L.E., Haraldsvik M. and Strøm B. (2008). Ressurser og Resultater i grunnopplæringen: Forprosjekt. *SØF-prosjekt nr.5100. Senter for Økonomisk Forskning*, Trondheim.

Bonesrønning, H. (2003) Class size effects on student achievement in Norway: patterns and explanations. *Southern Economic Journal 69(4)*, 952-965

Bonesrønning, H. (2004) The determinants of parental effort in education production: do parents respond to changes in class size? *Economics of Education Review 23*, 1-9

Bonesrønning, H. (2008) Peer group effects in education production: is it about congestion? *Journal of Socio-Economics*, 37: 328-342

Bonesrønning, H. (2011) Public Employees and Public Sector Reform Implementation. *Public Choice.* DOI 10.1007/s11127-011-9900-1

Bonesrønning, H and J.M.V Iversen (2008) Suksessfaktorer i grunnskolen: Analyse av nasjonale prøver 2007. *SØF Rapport 5/08. Senter for Økonomisk Forskning.* Trondheim

Bonesrønning, H. and Iversen J.M.V. (2011) Disadvantaged students in the early grades: Will smaller classes help them? *Education Economics, forthcoming.*

Bonesrønning, H., Iversen, J.M.V., and Pettersen I. (2010). Kommunal skolepolitikk etter kunnskapsløftet. *SØF Rapport 07/10. Senter for Økonomisk Forskning.* Trondheim

Bonesrønning, H. and Haraldsvik, M (2012) Peer effects on student achievement: Does the educational level of your classmates' parents matter? *Unpublished manuscript*

Borge, L.E., Haugsbakkken H., and Strøm B. (2009) Tidsbruk og Organisering i Grunnskolen: Sluttrapport. *SØF-rapport no 04/09. Senter for Økonomisk Forskning.* Trondheim

Bosworth, R., and F.Caliendo (2007) Educational production and teacher preferences. *Economics of Education Review 26,* 487-500

Carnoy, M and S. Loeb. 2003. "Does School Accountability Affect Student Outcomes? A Cross State Analyses." *Education Evaluation and Policy Analyses, 77, 337-358.*

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

Chetty, R., Friedman J.N., and Rockoff J.E. (2011) The long-term impact of teachers: teacher value-added and student outcomes in adulthood. *NBER Working paper #17699*

Christoffersen, H. and M. Paldam. (2003). Markets and municipalities: A study of the behaviour of the Danish municipalities. *Public Choice* 114: 79–102.

Couch, K.A. and D. W. Placzek (2010). Earning losses of displaced workers revisited *American Economic Review* 100 (1): 572-589.

Crockett, J., and Kauffman, J. (1999) The least restrictive environment: Its origins and interpretations in special education. *Mahwah, NJ: Lawrence Erlbaum Associates.*

Cullen, J.B. and R. Reback (2006). Tinkering Toward Accolades: School Gaming under a Performance Accountability System. *Advances in Applied Microeconomics* 14.

Dal Bo, E. 2006. Regulatory capture: a review. *Oxford Review of Economic Policy* 22: 203-225

Dee, T., and B. Jacob (2009) The impact of No Child Left Behind on student achievement. *NBER Working paper, no 15531*

Ermisch, J.F., and M. Francesconi (2001) Family structure and children's achievements. *Journal of Population Economics*, 14, 249-270

Figlio, D.N. (2007) Boys named Sue: Disruptive children and their peers. *Education Finance and Policy*, 2(4): 376-394

Figlio, D. and L. Getzler (2006) Accountability, Ability and Disability: Gaming the System? In *Advances in Microecenomics, Vol. 14: Improving School Accountability- Chekups or Choice?,* ed. T. Gronberg and D. Jansen, 35-49. Amsterdam: Elsevier.

Finn, J. D., and C. M. Achilles (1990) Answers and questions about class size: A statewide experiment. *American Educational Research Journal 27*, 557-577.

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

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

Hanushek, E.A. (1998) The evidence on class size. In Susan Meyer and Paul Peterson (ed.) *Earning and Learning: How Schools Matter?* Washington DC: Brookings Institution. 131-168

Hanushek, E. A., Kain, J.F., and Rivkin S.G. (1998). Does special education raise academic achievement for students with disabilities? *NBER Working Paper No 6690.*

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

Hanushek, E. A. and M. E. Raymond (2005) Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management 24(2)*, 297-327

Hoxby, C. (2000a). Peer effects in the Classroom: Learning from Gender- and Race Variations. *NBER Working Paper No. 7867.*

Hoxby, C. M. (2000b) The effects of class size on student achievement: New evidence from population variation. *The Quarterly Journal of Economics 115(4)*, 1239-1285.

Jacob, B. (2005) Accountability, Incentives and Behavior: Evidence from School Reform in Chicago. *Journal of Public Economics 89(5-6)*, 761-796.

Jacobsen, L.S., LaLonde R.J., and Sullivan D. G. (1993) Earning losses of Displaced Workers. *The American Economic Review* 83 (4): 685-709.

Jahnsen, H., Nergaard, S.E., and Flaatten S.V. (2006) I randsonen: Forekomst og Organisering av smågruppetiltak for elever på ungdomstrinnet som viser problematferd og lav skolemotivasjon. *Utdanningsdirektoratet*

Krueger, A. B. and D. Whitmore (2001) The effect of Attending a small class in the early grades on college-Test taking and middle school results: Evidence from project STAR, *The Economic Journal 111*, 1-28.

Kunnskapsdepartementet (2004) Kultur for læring. Available at http://www.regjeringen.no/nb/dep/kd/dok/regpubl/stmeld/20032004/stmeld-nr-030-2003-2004-.html?id=404433

Kunnskapsdepartementet (2010). Tid til Læring. Available at http://www.regjeringen.no/nb/dep/kd/dok/regpubl/stmeld/2009-2010/Meld-St-19-20092010.html?id=608020

Laffont, J., and J.Tirole, (1993). *A Theory of Incentives in Procurement and Regulation*, Cambridge, MA, MIT press

Lavy, V., and A. Schlosser (2011) Mechanisms and impacts of gender peer effects at school. *American Economic Journal: Applied Economics*, 3(2), 1-33

Lazear, E. P. (2001) Educational Production, *The Quarterly Journal of Economics 116(3)*, 777-803.

Leuven, E., Oosterbeek H. and Rønning M. (2008) Quasi-experimental estimates of the effect of class size on achievement in Norway. *Scandinavian Journal of Economics 110(4)*, 663-693.

Manski, C. (1993). "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies, Vol. 60, No. 3, 1993, pp. 531-542*.

Mckee, G., Rivkin, S.G., and Sims K.R.E. (2010). Disruption, Achievement and the Heterogeneous Benefits of Smaller Classes. *NBER Working Paper No. 15812*.

Nordahl, T. and Hausstätter R.S.(2009). Spesialundervisningens forutsetninger, innsatser og resultater. Høgskolen i Hedmark.

McCrary, J. (2008) Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics 142(2)*, 698-714

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

Nordahl, T. og R.S. Hausstätter (2009) Spesialundervisningens forutsetninger, innsatser og resultater. Hedmark University College.

OECD (2012) OECD Family Database, OECD, Paris (www.oecd.org/social/family/database)

Painter, G. og D.I. Levine (2000) Family structure and youths' outcomes: which correlations are causal? *Journal of Human Resources* 35, 524-549

Public Agenda (2004). Education : Teaching Interrupted: Do Discipline Policies in Today's Public Schools Foster the Common Good? Accessed at: http://www.publicagenda.org/research/research_reports_details.cfm?list=3

Rattsø, J., and R.J. Sørensen, (2004). Public employees as swing voters: Empirical evidence on opposition to public reform. *Public Choice* 119: 281–310.

Sørensen, R., and A.H. Bay, (2002). Competitive tendering in the welfare state: perceptions and preferences among local politicians. *Scandinavian Political Studies* 2: 357-384

The Norwegian Institute of Public Health (2012): Facts about ADHD http://www.fhi.no/eway/default.aspx?pid=233&trg=MainLeft_5648&MainArea_5661=5648:0:15,2917:1:0:0:::0:0&MainLeft_5648=5544:60586::1:5647:2:::0:0

Tirole, J. (1986) Hierarchies and Bureaucracies: On the Role of Collusion in Organizations. Journal of Law, *Economics and Organization*, 2(2): 181-214

Urquiola, M. and E. Verhoogen (2007) Class size and sorting in market equilibrium: theory and evidence. *NBER Working paper No. 13303*

Utdanningsdirektoratet (2009) Veileder til opplæringsloven om spesialpedagogisk hjelp og spesialundervisning. Available at www.udir.no

Woessmann, L. (2005) Educational production in Europe. *Economic Policy 20(43)*, 445-504