



AARHUS UNIVERSITY



# Coversheet

---

**This is the accepted manuscript (post-print version) of the article.**

Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

**How to cite this publication**

Please cite the final published version:

Junge, S. Y. (2022). Administrative Groupings and Equality in Public Service Provision. *Journal of Public Administration Research and Theory*, 32(2), 252–268. <https://doi.org/10.1093/jopart/muab031>

## Publication metadata

**Title:** Administrative Groupings and Equality in Public Service Provision  
**Author(s):** Sarah Yde Junge  
**Journal:** *Journal of Public Administration Research and Theory*, 32(2), 252–268  
**DOI/Link:** <https://doi.org/10.1093/jopart/muab031>  
**Document version:** Accepted manuscript (post-print)

**General Rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Administrative Groupings and Equality in Public Service Provision

Sarah Yde Junge

Aarhus University <sup>\*†‡</sup>

July 2021

## Abstract

Scarcity in public service agencies requires a prioritization of resources, and inherent to all prioritizations is a comparison of the cases. Despite the amount of research that has been conducted on the prioritization process, surprisingly little attention has been paid to the underlying comparison. This study suggests that the administrative grouping of citizens in public service agencies influences the comparison so that the allocation of public services not only depends on a citizen's need, but also on the needs of the other citizens assigned to the same agency. Based on registry data on more than 300,000 students, this study exploits the as-good-as-random assignment of students to classes within schools to analyze the influence of classroom composition on referrals to special education in Danish primary school. The result

---

\*Department of Political Science, Aarhus University, Aarhus C, Denmark

†TrygFonden's Centre for Child Research, Department of Economics and Business Economics, Aarhus University, Aarhus, Denmark

‡Centre for Integrated Register-based Research, CIRRAU, Aarhus University, Aarhus, Denmark

shows that students with better-performing peers are considerably more likely to be referred to special education than equally low-achieving students in lower-performing classes. This finding illustrates an unwarranted distributive consequence of the administrative grouping of citizens, and thereby initiates a discussion of how groupings affect service provision and how such an influence can be mitigated.

## Introduction

Citizens receiving public services are almost always grouped in some way in order for the bureaucracy to be able to provide them the service. Examples are students in schools and classes, toddlers in preschools, patients in hospital wards, prisoners in blocks, the assignment of patients to a general practitioner, the group of clients a specific caseworker handles, the assignment of litigation to judges or the division of neighborhoods in police districts. These administrative groupings of citizens may seem simply to be an organizational arrangement of cases, but the question is whether the grouping in itself also influences who gets which services, when and how. This question is the center of this study. The provision of public services is very often based on street-level bureaucrats' discretion in their interactions with citizens (Lipsky, 1980). There is a large body of literature studying how this discretion may lead to differential treatment of citizens in street-level bureaucracies (e.g. S. C. Andersen and Guul, 2019; Maynard-Moody and Musheno, 2003; Soss, Fording, and Schram, 2011 and see Tummers et al., 2015 for a review). Many of these studies implicitly assume that street-level bureaucrats compare their clients in order to figure out whom to prioritize. Focusing on this prioritization, this study theorizes about and investigates how the administrative grouping of citizens may be an overlooked source of differential treatment in public service provision. Despite the amount of work that has been done on street-level bureaucracy and despite the obvious unfairness of the administrative grouping of citizens as a determinant of the service

level provided to each individual citizen, the consequences of the grouping itself have not yet received much attention in the public administration literature.

The aim of this study is first to conceptualize administrative grouping, then to combine different theoretical insights in a theory of how the administrative grouping of citizens in public services affects the allocation of services, and lastly to make a first attempt at testing this theory. Legal regulation as well as the bureaucratic organization of public service agencies ought to secure equal treatment of citizens (Weber, 1947). However, theoretically, one could imagine at least three mechanisms in which the grouping of citizens could in fact influence the service provided to each individual citizen. First, the other group members may claim a smaller or larger part of the available resources and thereby influence what is left for any individual group member. Second, when street-level bureaucrats assess each citizen's need for a given service, the needs of other group members may be used as a baseline for comparison (Lipsky, 1980; Tversky & Kahneman, 1974). Last, the group composition may signal something to the individual group member about her own needs and abilities that could make her change her demands for public service (Festinger, 1954). All three mechanisms have the same implication, namely that citizens grouped with high-needs citizens are less likely to receive public services than citizens in a low-needs grouping. While the first mechanism may be a legitimate result of different political prioritizations, the other two mechanisms are not. This study keeps the resources fixed, and investigates whether citizens' eligibility for public services depends on the needs of the other citizens in the same administrative groupings within public service agencies.

Administrative groups may be difficult or unethical to experimentally manipulate, and non-manipulated groups are very often non-randomly created, which may cause studies of these groupings to be biased. Therefore, the primary concern in studies of administrative groupings is to overcome selection issues in a non-experimental setting. To explore the influence of grouping and at the same time overcome selection issues, this study relies on

the grouping of students in classrooms within schools. The school system represents a large and important part of public service provision in terms of tasks, finances and personnel, and schools are decentralized units with considerable room for discretion in their service provision (Meier & O’Toole, 2006). Referrals to special education, like many public services, are expensive and provided based on individual assessments. This case thereby provides a good starting point for a test of the theoretical argument about the influence of grouping on public service provision. A fixed effects design enables me to deal with selection into schools and uneven allocation of resources between schools. Formal tests of assignment to classes show that classroom formation within schools is balanced on all important observable variables, and thereby provides as-good-as-random variation in the groups. I use this variation to investigate how the administrative grouping of students affects referrals to special education in Danish public primary school. The Danish registers contain population-wide data ( $N = 314,405$ ) on special education referrals, classroom composition and academic performance in national standardized tests, as well as a comprehensive list of background information. Using this, I can investigate how variations between classes within schools may affect the probability of an individual being referred to special education (Ammermueller & Pischke, 2009; Hoxby, 2000).

In line with the theoretical claim, the results show that students with better-performing peers are more likely to receive special education. This study draws attention to a new source of inequality in public service provision, which has a number of implications for both public administration research and the organization of public service provision. First, it builds a theory about how administrative groupings may cause inequality and how this inequality can be mitigated. Second, it calls for further research into the influence of administrative grouping and the general relevance of this phenomenon across different street-level bureaucracies. Finally, an awareness of inequalities as a result of administrative groupings of citizens may inform the processes of decentralization and organizational design of street-level bureaucracies

to increase the likelihood of public services being allocated fairly.

## Theory

The overarching theoretical aim of this study is conceptualize administrative grouping and combine different theoretical insights in a theory about how the administrative grouping of citizens affects the provision of public services in street-level bureaucracies. The theoretical argument for a relation between these originates from the street-level bureaucracy literature. Below, I will review the literature on street-level bureaucracy, elaborate on the theoretical concept of administrative grouping and combine this with insights from related literature in a theory about how the administrative grouping of citizens influences public service provision.

### Street-level bureaucracy

Weber argues that a fundamental advantage of modern bureaucracy is that "Everyone is subject to formal equality of treatment, that is, everyone in the same empirical situation" (Weber, 1947, p. 340). Such an equality necessitates tight regulation of the bureaucracy from the politicians to the street-level bureaucrat, where the implementation of services and sanctions follows a detailed set of rules and guidelines. According to Weber, the perfect bureaucracy is "dehumanized," and the processing of cases is completely objective and independent of "personal, irrational and emotional elements which escape calculation" (Weber, 1922 [1968], p. 975). However, reality is often so multifaceted and complicated that it is impossible to regulate every situation in a way that allows street-level bureaucrats to just objectively rely on rules and guidelines in the implementation of policy. Contrary to Weber's bureaucracy ideal, street-level bureaucrats often have substantial discretion in the determination of citizens' rights, access, eligibility and the imposition of sanctions (Brodkin, 1997; Lipsky, 1980; Meier & O'Toole, 2006). Exactly for this reason, Lipsky (1980) argues that

street-level bureaucrats occupy a critical position in society. When street-level bureaucrats, such as teachers, social workers, judges or police officers, assess citizens' rights, needs, access and eligibility, they decide how to implement policies, and thereby become day-to-day policymakers. Through their discretionary decisions, street-level bureaucrats have influence over significant aspects of citizens' lives (Lipsky, 1980; Maynard-Moody & Musheno, 2003; Olson, 2016). This discretion is to some extent an inevitable condition of street-level bureaucrats' jobs, and it enables them to use their expertise and to put their professionalism into play when dealing with complex cases (Lipsky, 1980; Thomann, van Engen, & Tummers, 2018). On the other hand, the lack of political control over the implementation process may entail the risk of an implementation that does not fulfill its political objective or the promise of modern bureaucracy to treat every citizen in the same empirical situation equally (Bawn, 1995; Huber & Shipan, 2002; Lipsky, 1980).

Street-level bureaucrats often experience high workloads, which may lead them to exercise their discretion in ways that enable them to cope with this pressure. Examples of coping include bending or breaking rules to help clients, but also prioritizing some clients over others and routinizing or rationing services (see Tummers et al., 2015 for a review). Though active coping behavior is often intended to help the citizens, the pressure may also increase the risk of street-level bureaucrats relying on stereotypes or perceptions of worthiness, which may cause differential treatment (examples are discrimination based on race (S. C. Andersen and Guul, 2019; Olson, 2016; Pedersen, Stritch, and Thuesen, 2018), gender (Lavy and Sand, 2018; Wenger and Wilkins, 2008), characteristics of the street-level bureaucrat (Jensen and Pedersen, 2017) and perceptions of the citizen (Guul, Pedersen, and Petersen, 2021; Maynard-Moody and Musheno, 2003; Sandfort, 2000; Tummers, 2017)). Organizational factors like professional norms, management, deliberation, political control and accountability schemes have been shown to shape the decision-making processes of street-level bureaucrats and the level of coping (e.g L. B. Andersen, 2009; Brodtkin, 1997, 2011; Chaney and Saltzstein, 1998;

Jacob, 2005; Møller, 2021; Tummers et al., 2015).

This study investigates how another organizational dimension, namely the administrative grouping of citizens, may affect street-level bureaucrats' discretion. Below, I will elaborate the theoretical concept of administrative grouping and highlight two organizational characteristics on which the groupings may differ between agencies.

## **Administrative grouping**

The division of labor in public services necessitates some kind of assignment of citizens to street-level bureaucrats. This assignment creates administrative groups of citizens for which a given street-level bureaucrat or a group of street-level bureaucrats are responsible in their work. Examples of such administrative groups could be the group of patients assigned to a general practitioner, the group of students assigned to a specific classroom or the assignment of litigation to a judge. The allocation of citizens to and within public services constitutes the administrative grouping. The groups may be identifiable on different levels of the street-level bureaucracies, as clients of a specific public agency, and also as parts of smaller units within the agency. Regardless of the organizational level, all administrative groupings are administratively created and known by the public service agency.

Given the diversity of public service agencies, the organizational characteristics of administrative groupings may also differ. Here I will highlight two dimensions on which groupings may differ, and which may affect the importance of the group for the service provision. First, the groupings may be created based on different organizational specializations, and second, the citizens assigned to the same group may or may not interact with each other. Simon (1945) distinguishes between specialization by place and specialization by function. Street-level bureaucracies that specialize by place are spatially divided in a number of parallel agencies that provide the same service, e.g. schools, general practitioners and nursing homes, whereas agencies specialized by function create administrative groups based on the service



they provide, such as hospital wards, which are specialized in different illnesses and treatments, or prisons, whose security standards match the severity of the prisoners' criminal acts. The organizational specialization guiding the administrative grouping at the agency level may also entail other differences. When administrative groupings are based on specialization by functions, the groupings may be more homogeneous with regard to the public service provided, and the street-level bureaucrats working with these groups may become experts in their specific fields to a larger extent than street-level bureaucrats who work with spatially defined groups. Furthermore, when services are spatially divided in parallel agencies, citizens will often have a larger say in which group they are assigned to. In schools, daycares and nursing homes, for example, citizens often hold some knowledge on the difference between institutions and they may also have quite substantial influence over which institution they are assigned to, and hence which group they become part of—if not directly through their choice of service provider, then by moving to another district (Tiebout, 1956).

The other dimension concerns the interaction between group members. While some services are delivered within the administrative groups, other services are delivered individually, without any interaction between citizens assigned to the same administrative group. This characteristic may affect the relevance of the group for citizens as well as street-level bureaucrats. If the citizens do not know their group composition, it is unlikely that the administrative group would cause a change in their behavior towards the agency. The grouping is by definition known to the street-level bureaucrat, but one could imagine that the grouping would be more present in the mind of the bureaucrat when the group members are observed at the same time.

The administrative groupings of citizens may become influential when street-level bureaucrats have discretion in their work, because this discretion may create room for consideration of characteristics not strictly related to the eligibility for services. The administrative grouping may provide citizens as well as street-level bureaucrats with a comparative standard of

needs and eligibility that may inform decisions regarding service provision. Below, I will elaborate on three different mechanisms through which the comparative standard provided by the administrative grouping may influence the provision of public services.

## **Arguments for an influence of grouping**

When service provision is not automated but relies on street-level bureaucrats' assessments, the administrative group may affect the service provided for each citizen through at least three different mechanisms. The first mechanism concerns the limited resources available in a street-level agency. Because there is no price limit on the demand for public services, it very often exceeds the supply and necessitates a prioritization of cases (Lipsky, 1980). If agencies prioritize their resources to the citizens who need them most, this will cause citizens in low-need groups to experience a higher service level than equally deserving citizens in a high-need grouping. Such differences can be eliminated politically with financial equalization of the differences in demand and supply between agencies (Boadway & Shah, 2009). Decisions of financial support and equalization are politically determined, and therefore differences between agencies must be considered to have been politically accepted. By contrast, differential treatment based on the administrative groupings of citizens *within* agencies violates the promise of modern bureaucracy. Therefore, the remaining part of this study will discard the prioritization mechanism and focus on the two other mechanisms for the influence of administrative grouping in the provision of services to citizens within the same agency.

Within agencies, everyone draws from the same pool of resources, and therefore differential treatment of similar cases is not caused by differences in the available resources. Below, I will outline two theoretical mechanisms through which administrative groupings may be a source of differential treatment of empirically similar cases. The first mechanism is a supply-side mechanism concerned with how the street-level bureaucrat makes assessments of needs and eligibility, and the second mechanism is a demand-side argument concerned with citizens'

perceptions of their own needs. They rely on the same logic about the administrative group as a relevant comparison for assessments of needs, and will not be separated in the empirical analysis of this study. Instead they are used as a theoretical foundation to understand why we could expect administratively created groups to influence the services provided for citizens in public agencies, beyond the politically accepted differences caused by differential allocation of resources between agencies.

The supply-side mechanism of an influence from the administrative grouping is that the grouping may affect the assessments street-level bureaucrats make. An argument for such a correlation is found in psychological work on human processing of information and construction of judgment, which has repeatedly demonstrated that humans cannot make absolute judgments. Instead, they use existing knowledge or prior experiences to anchor their assessments. Thus, human evaluations always rely on some kind of comparison (Helson, 1947; Kahneman, 1992; Tversky & Kahneman, 1974).

In public administration research it has been demonstrated that managers use prior performance and the performance of neighboring organizations as reference points that inform expectations and aspiration levels (Holm, 2017; Nielsen, 2014; Simon, 1939), and other studies have investigated the extent to which citizens also rely on comparisons when they evaluate performance information on public services (S. C. Andersen & Hjortskov, 2016; Olsen, 2017). However, whether street-level workers compare their clients to determine eligibility is still unknown. In his work on street-level bureaucracy, Lipsky mentioned in passing that *"it is probably fair to say that clients will always be differentiated in terms of their perceived relative normality, regardless of how absolutely receptive to intervention they are"* (Lipsky, 1980 [2010], p. 113). Thus, Lipsky suggested that street-level bureaucrats also rely on some kind of comparison to other clients when they evaluate clients' needs. This comparative element is also an underlying assumption in much work on street-level bureaucrats' coping. One example is *creaming* or *prioritizing*. To prioritize the easiest clients, street-level bureaucrats

need to compare their clients to each other to figure out who they can most easily succeed with, and thereby who they should care most about in order to reach their performance goals. Hence, the argument is that the administratively created group of citizens that a street-level bureaucrat is working with affects her perception of what is normal, who is in need of public services, and who is considered eligible for them.

Some empirical studies from related fields support the notion of a group-dependent bias in street-level bureaucrats' assessments of citizens. For example, studies of grading and teachers' recommendations of further education for their students suggest that teachers use the classroom as a frame of reference, which implies that students get better grades (Calsamiglia & Loviglio, 2019; Møen & Tjelta, 2010; Neumann, Trautwein, & Nagy, 2011), and are more likely to get better recommendations for further education, when their peers do worse (Boone et al., 2018). However, most of these studies do not methodologically address the selection into groups, and they all study a quite uni-dimensional assessment, with no economic costs associated. Public services are often based on more complex assessments, and the decision to provide a service for a citizen is associated with a cost. The studies above inspire the argument that administrative groupings may affect service provision, but so far there is no empirical evidence that their conclusions will withstand a harder test and travel to more complex cases like provision of services.

The demand-side mechanism instead points to the citizens. The argument is that the grouping itself may signal something to group members about their own status, which makes them change their demands for public service. This argument emanates from social comparison theory, first introduced by Festinger (1954). The basic claim is that humans rely on reference groups when they evaluate their own abilities and opinions, and may think more or less of themselves depending on the reference group. Empirical studies have demonstrated the importance of the reference group in assessments of academic self-efficacy, showing that students evaluate their own performance more favorably and have higher aspirations when

they perform relatively better than their peers (Davis, 1966; Marsh, 1984; Marsh et al., 2008; Seaton, Marsh, & Craven, 2010). These studies indicate that humans' assessments of themselves adapt to their surroundings. One could imagine that this adaptation would also occur in public services where members of administrative groupings interact with each other, such that citizens grouped with others who have low needs would think more of their own needs and hence increase their demands for public support.

Together these two mechanisms form the hypothesis that *a citizen in an administrative group where the other group members have lower service needs is more likely to receive public services than the same citizen would be in an administrative group with higher service needs.*

The demand-side mechanism and the supply-side mechanism work in the same direction, and therefore each of them alone could cause the administrative grouping to produce differential treatment. While I will not separate the two mechanisms in the analysis below, it is probably fair to assume that the demand-side and the supply-side mechanisms also interact. When the street-level bureaucrats' assessments are influenced by the composition of the administrative group, citizens with high needs compared to the remaining group may feel this in their interaction with the street-level bureaucrat, which could amplify their own perception of needs. Likewise, service demands from citizens with comparatively high needs may amplify a group bias held by the street-level bureaucrat because they increase the salience of the group composition and because these demands may increase the bureaucrat's workload, which could cause an increased reliance on the composition cue received from the administrative grouping.

## Empirical design

There is one major challenge to studying the consequences of administrative groupings—namely to overcome the selection issues. Even when citizens do not decide who they will be grouped with, administrative groups are seldom created completely at random, and therefore the con-

sequences of groupings are hard to separate from the constituent reasons. It is rarely possible for researchers to experimentally manipulate groupings, and the complexity of a group may be hard to convincingly simulate in a survey-experimental design. Therefore, the study of groupings needs to rely on observational data with some kind of exogenous variation in the group formation.

I use register data on special education referrals in Danish primary schools to examine the consequences of administrative groupings for service provision. The grouping of students in schools and classes provides a clearly defined and rather constant measure of grouping, which eases an analysis of the influence of grouping. Formal tests of differences between classes on predetermined student characteristics show that assignment to classes within schools is as good as random. These differences between classes within schools provide exogenous variation to the administrative groups under study, which enables me to overcome selection issues. Furthermore, I rely on school and register data, which both provides a sample large enough to detect even smaller effects and also ensures very high data validity. The educational system provides a good starting point for the study of grouping effects for multiple reasons. The school sector represents a large and important part of public service provision and schools are highly professional and decentralized agencies, which leaves substantial room for teacher discretion (Meier & O'Toole, 2006). The evaluation and assessment of students' abilities, needs and well-being is a core task for teachers, and the decisions teachers make are often of great importance to the student. Furthermore, decisions of special education referral, like the provision of many public services, are associated with a non-negligible cost. Though there are some organizational structures that apply specifically to schools, such as the high degree of interaction among citizens within groups, the stability of the administrative groupings over time and the substantial self-selection into schools, the case of student referrals to special education captures the discretion and the complexity as well as the individual stakes of many decisions of public service provision in modern welfare states. The school fixed effects design

overcomes the selection issue, and I will return to the other issues and what we can learn about administrative grouping in general from this case in the discussion section.

Below, I introduce the case of special education in the Danish school system and the estimation strategy applied, followed by an operationalization of the measures used in the study and a discussion about the assumption of random assignment to classes.

## **Special education referrals in Danish schools**

In Denmark, there is a voucher system that allows Danish parents to choose which school to send their children to. There is thus considerable selection into schools, but the assignment to classes within schools does not seem to suffer from this selection. Assignment to classes is done at the school level. National regulations of class formations state that there can be a maximum of 30 students per class and that the local governments must make guidelines for class formation (“Folkeskoleloven”, 2017). Most municipalities only make guidelines about class sizes, and let the schools themselves decide how to assign students to classes. A look into different school policies on class formation shows that most schools aim for a balanced gender distribution, and some also actively assign students to classes in ways that secure a social and geographic balance between classes. In some schools, parents may name a child whom they want their child to be assigned to class with. The school principal assigns teachers to classes. Each class has one or two head teachers, who typically teach Danish and math. Head teachers are normally assigned to a class for a 3-5 year period, so the head teachers in 1<sup>st</sup> grade will follow the class until 3<sup>rd</sup> to 5<sup>th</sup> grade.

When there are students in a class whose "development necessitates special consideration or support that cannot be given within the setting of the ordinary class" (“Specialundervisningsloven”, 2010), they are eligible for special education. Students can be referred to special education for many different reasons, but the purpose of special education is for the students to acquire knowledge and competencies that will prepare them for further education. There-

fore, academic challenges need to be present no matter what other difficulties the student may struggle with (“Folkeskoleloven”, 2017; “Specialundervisningsloven”, 2010).

Detailed student-level data on special needs are rarely available, and especially not for the entire population of students, but in the school year 2011-12, special education referrals and the magnitude of special education was registered for every student in the Danish school system. Performance data on an individual level is available from 2<sup>nd</sup> grade for all public school students, and therefore I can include everyone in ordinary classes in public schools in grades 3 to 9 in the school year 2011-12 in the analyses. This gives me a sample of 314,405 students. This unique data source is the basis of this study.

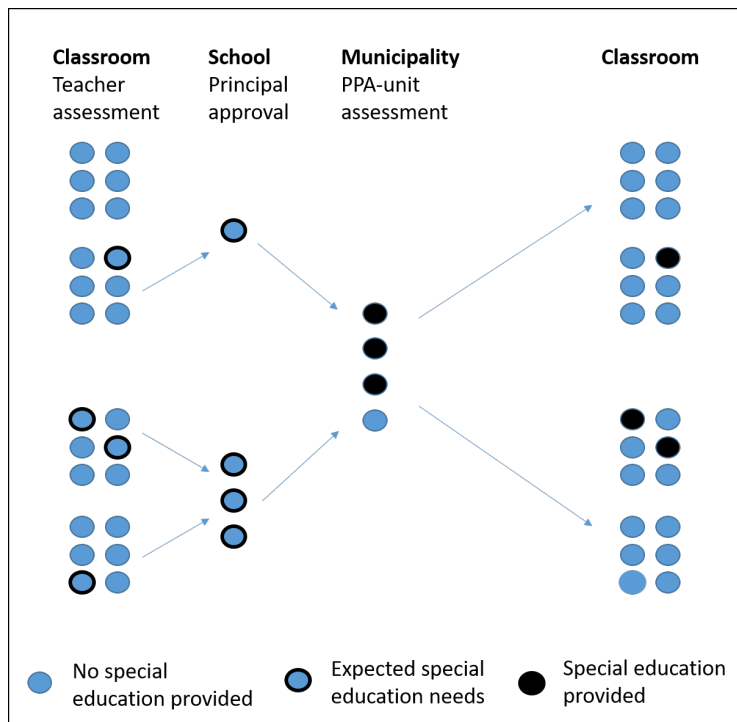
While special education can be organized in several different ways, the most important distinction for this study is whether the student with special needs is enrolled in an ordinary class or in a special needs class. Only students in ordinary classes are included in this study. Students in this category will attend an ordinary class most or all of the time and then have either an extra teacher to support them during some of the lessons or have some individually scheduled lessons in addition to or instead of ordinary class. Even though there is some debate about the effect of special education (Hanushek, Kain, & Rivkin, 2002; Keslair, Maurin, & McNally, 2012; Kvande et al., 2019), there is no doubt that special education is given to students in an attempt to support them. This is underlined by the amount of money spent on special education. A report on special education in Danish primary schools estimates that the average cost of a student referred to ordinary special education is 168 % of the cost of an ordinary student (Ministry of Finance, 2010).

Figure 1 illustrates the referral process. When teachers feel that a student might benefit from special education, they are obliged to recommend the student for it. The recommendation is received by the headmaster or a committee at the school that takes care of special education. From here, the standard procedure is that the headmaster or the committee (in consultation with the parents) will ask a centrally organized pedagogical-psychological advi-



sory unit (PPA unit) to make an assessment of the student. If the PPA unit finds that the student should be granted special education, it will also make a recommendation about the type and amount of special education the student should be granted. Thus, the teachers do not decide who will ultimately receive special education or how much; but if the teachers are not aware of a student's challenges or needs, the process will not be initiated. Hence, the teachers are important for the decision of who will receive special education, whereas the PPA units determine both who will receive special education and how much. When a student has been referred to special education, her needs and the type of support granted have to be reconsidered at least once a year.

Figure 1: The referral process



The frequency of special education referrals differs greatly between Danish municipalities. A governmental report from 2010 stated that "It is assessed that the variations between the municipal referral patterns are an expression of the fact that there is no clear and uniform practice across the municipalities as to what types and degrees of issues are dealt with in

ordinary special education" (Ministry of Finance, 2010). The lack of a clear and uniform practice implies that special education referral is a service that is provided to students on a heavily discretionary basis.

## Estimation strategy

As argued above, a simple comparison between schools would not necessarily shed light on the impact of groupings in schools, because the selection into schools may cause the student populations to differ between schools. To overcome these selection issues, this study applies a fixed effects design. Inspired by Hoxby's (2000) exploitation of idiosyncratic cohort-to-cohort variations in her study of peer effects, I examine whether class-to-class differences within the same school are systematically associated with referrals to special education.

The main specification of the model applied to investigate the influence of the group on special education referrals is:

$$y_{ics} = \alpha_s + \beta \bar{X}_{cs} + \gamma X_{ics} + \delta \mathbf{Z}_{ics} + \eta \bar{\mathbf{Z}}_{cs-i} + C_{cs} + \mu_{cs} + \epsilon_{ics}^1,$$

where  $y_{ics}$  is an indicator of whether student  $i$  in class  $c$  in school  $s$  is referred to special education.  $\bar{X}_{cs}$  is the grouping effect that is the average class need for special education.  $X_{ics}$  is the student's individual need for special education.  $\alpha_s$  is the school fixed effects.  $\mathbf{Z}_{ics}$  is a vector of individual background characteristics,  $\bar{\mathbf{Z}}_{cs-i}$  is a vector of leave-me-out class averages on background characteristics and  $C_{cs}$  is an indicator of the grade level.  $\mu_{cs}$  is the class-level error term and  $\epsilon_{ics}$  is the individual-level error term.  $\mu_{cs}$  represents the correlated effect, which is the unmodeled class-specific influences. If  $\mu_{cs}$  is correlated with the class composition, this will cause a bias in the estimates. If students and teachers are randomly

---

<sup>1</sup>In line with the theoretical argument about the teacher's general impression of the class, the analyses use the simple average of the class' needs and the more appropriate leave-me-out average of the background characteristics. The alternative specification with a leave-me-out average of the class' needs does not change the results significantly.

assigned to classrooms, the correlated effect will not bias the model. The school fixed effect  $\alpha_s$  removes any selection effect caused by selection into schools by removing all variation between schools, including any differences in available resources.

As a supplement to the analysis of referrals, I conduct an analysis with the amount of special education each student receives as the dependent variable. As described above, the amount of special education is externally determined, and hence the classroom is not expected to influence the amount of special education that referred students receive. If the amount of special education is correlated with the composition of the class, the classroom measure may be correlated with critical omitted variables. If the amount of special education is in fact independent of the classroom composition, a dependency between classroom composition and referrals are very likely a result of the influence from the administrative grouping of students.

Given the dichotomous nature of the dependent variable, I estimate the effect of classroom composition using a linear probability model. The linear probability model has the advantage of being more efficient and easier to interpret than its nonlinear counterparts (Angrist & Pischke, 2008, p. 107), especially in hierarchical models. Robust standard errors are applied in all specifications of the model to account for the heteroskedasticity imposed by the OLS estimator when estimating models with a binary dependent variable. The main specification is also estimated using a constrained conditional logistic regression. Following Beck (2020), the results from the logistic regression as well as the results from a linear regression using the logistic sample are available in the appendix. The effect sizes from these estimations are substantial and statistically significant, like the results from the main specification presented below.

## Measures

A crucial part of the concept of administrative groupings is to define which characteristics of the group may influence the service provided for each individual group member. The

influential characteristics are probably a multi-faceted pool; however, given the proposed mechanisms above, I will argue that the relevant characteristics of other group members need to be the characteristics that determine their needs or eligibility for public services. Special education referrals are complex decisions and one may assume that many different factors could influence students' needs for special education. However, since academic performance is of central concern to decisions about special education, this study uses the academic performance of the class as the group measure, despite the fact that actual assignment to special education probably relies on more nuanced assessment.

I rely on the Danish national tests in reading to measure the academic ability of the individual student and the class. The national tests are computer-based, adaptive and self-correcting (Beuchert & Nandrup, 2014). This means that human bias in their correction is minimized and, even more important, the students' teacher has no influence on the scoring of the test. The tests are mandatory for all public school students in 2<sup>nd</sup> grade, 4<sup>th</sup> grade, 6<sup>th</sup> grade and 8<sup>th</sup> grade.

The national tests are taken in the spring, but special education is registered in October. To make sure that the test scores are measured prior to special education referral, I use the test scores achieved in the previous year for students in grades 3, 5, 7 and 9, and two years earlier for students in grades 4, 6 and 8. The test scores are standardized within grade and year on the population of test scores. The class measure is the average test score performance of the class. It seems unlikely that teachers actually calculate this, but the teachers are assumed to have an impression of the general academic level in the classroom, and I expect the average performance of the class to mirror this impression.

The dependent variable of the study is the dichotomous variable indicating whether a student is referred to special education. It takes the value 1 if the student is referred to special education and 0 otherwise. 7.23 % of the students in the sample were referred to special education. An alternative specification of the model uses the hours of special education a

student receives per week. For students not referred to special education, this variable takes the value 0. The information on the amount of special education is known for 77.3 % of the referred students. On average they received 1 hour and 36 minutes per week, though the average amount of special education received by special needs students may be lower since the 22.7 % whose amount of special education is unknown performed slightly better academically (mean -0.77, sd 1.03) than the 77.3 % who were observed (mean -0.86, sd 0.99)

A set of background variables obtained from the Danish registries is included in the analyses. These are gender (0=girl, 1=boy), immigrant status (0=Danish, 1=immigrant or descendant of immigrants), psychiatric diagnoses (0=no diagnosis, 1=diagnosis) and average parental education in years. In cases where only one parent’s education is known, the measure relies on this parent alone. These measures are included as control variables at the individual level as well as classroom leave-me-out averages. As a measure of resource allocation, class size as registered at the beginning of the school year is included in a supplementary analysis.

Table 1: Descriptive statistics

	Mean	SD
Share referred to special education	0.072	
Individual test score	0.041	0.959
Parental education	14.220	2.091
Share of boys	0.502	
Share of immigrants	0.097	
Share with psychiatric diagnosis	0.066	
Class size	21.461	3.596
Classes per school cohort	3.178	1.139
Number of schools	1,330	
Number of classes	15,705	

Because the administrative groups are a crucial component of this analysis, one needs to think carefully about how missing data on the individual level may affect the group measures. To be sure of the reliability of the measures of class composition, classes where more than three students have not taken the test are excluded. Analyses of the subsample of cohorts

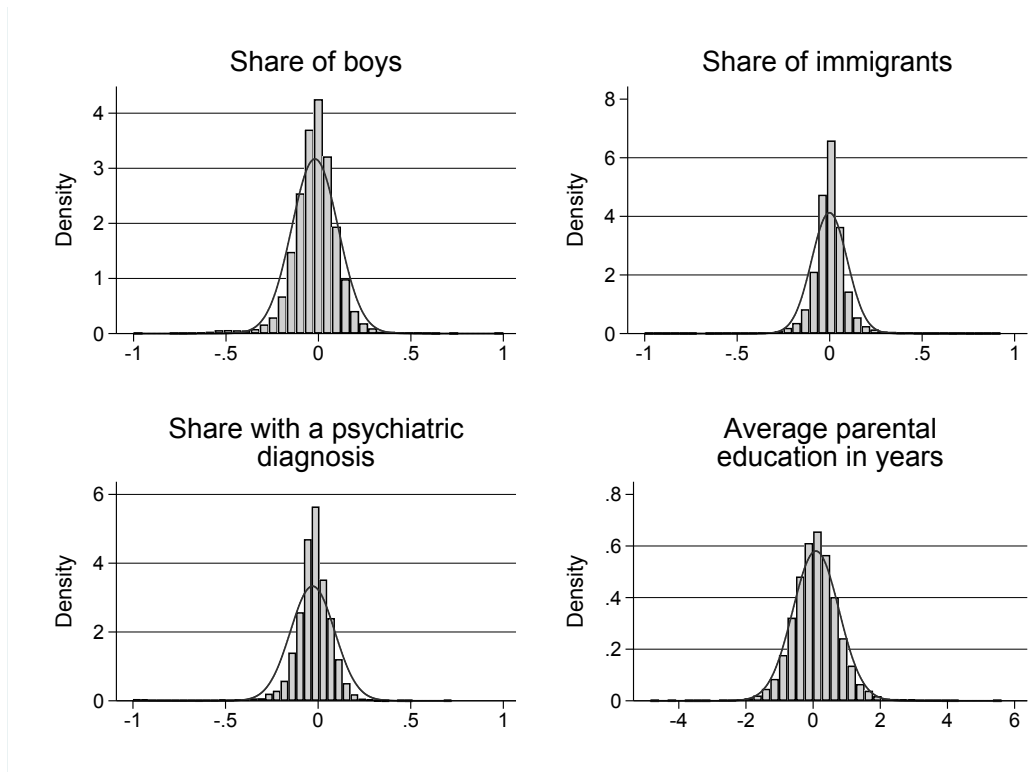
with no missing test scores (see Table A1 in the appendix) reveal an effect size larger than the one reported here. This implies that the estimates reported in the results section are not driven by uncertainties of the true class average. Furthermore, observations in classes with fewer than five students or more than 30 students are excluded, since these classes are likely to be a result of an error in the registration process rather than the real class size. All class-level variables are standardized within grade based on the sample of classes used in this study.

## Assignment to classes

One might worry that parents not only select into schools, but also into classes within schools. There is nothing in the guidelines on class formation that indicates that this is the case, but they do not exclude the possibility that parents may interfere in class formation. To be sure class formation is not subject to parental selection, I conduct a series of tests below to formally test the assumption of random assignment to classes within schools. Figure 2 shows histograms of the deviation from the classroom average to the average of the remaining cohort within a given school on predetermined measures. If differences between classes are random, the deviations pictured in Figure 2 would be normally distributed. As can be seen in Figure 2, the deviations for all variables follow the plotted normal distribution curve, though the differences seem to be a bit smaller than one would expect from completely random assignment. Following the procedure used by Ammermueller and Pischke (2009), I also make a statistical test of whether students are assigned to classes in any systematic way using a Pearson  $\chi^2$  test for observable characteristics that could potentially cause selection bias. For parental education, which is a continuous variable, I use the Kruskal-Wallis test. The p-values from these tests are plotted in Figure A1 in the appendix.

The principles of class formation imply that the patterns revealed in Figure 2 may not be an expression of true random assignment, but rather a deliberately balanced assignment,

Figure 2: Classroom deviation from school-cohort average



at least for some of the variables. However, since the allocation is balanced, the variation in peer composition across classes can be considered as good as random for statistical purposes (Calsamiglia & Loviglio, 2019). Furthermore, I include the background characteristics as control variables in the models, to account for any non-randomness in these. Taken together, the tests above indicate that the underlying assumption about as-good-as-random assignment to classes within schools is not violated, and hence that the fixed effects approach manages to overcome selection issues. Another concern from  $\mu_{CS}$  is that resources and teachers may be assigned to classes dependent on the classroom composition. This concern is harder to deal with, because there is no information on teacher assignment available. Ammermueller and Pischke (2009) argue that "teacher-shopping"—that is, teachers changing classes due to the ability of the students—is not very common in European school systems, which would imply that the teacher assignment is not problematic in this case. To try to overcome the resource

question, I conduct three robustness checks. First, I investigate whether a smaller class size is correlated with fewer referrals. If the classes are smaller, one would expect the teacher to have more time within the ordinary class to support students with special needs. Second, I limit the analysis to 3<sup>rd</sup> grade students, the earliest year for which data is available. Teachers in Danish primary school are typically assigned to classes for a three- to five-year period, so most students will have the same teacher in Danish from grades 1 through 3. When the students start, the school may not know which of the classes will benefit the most from a better teacher, and hence the assignment of teachers in the lower grades may be considered more random than in higher grades. Third, I further limit the 3<sup>rd</sup> grade sample to students who do not have older siblings and do not live in a home with older children. This limitation excludes students with parents who have other children at the school, and hence possess some knowledge of the teachers and their qualifications, which they may use to affect the assignment of teachers to classes. The results of the robustness checks are shown alongside the results of the analysis in the results section.

A final concern to be mentioned here is whether there is sufficient variation between classes within schools to estimate the effect of classroom composition on special education referrals. A variance decomposition of class-level averages of reading scores is shown in Table 2. Not surprisingly, the majority of the variation in class-level reading scores is between schools (around 70 %). However, there is still substantial variation in class-level averages left within schools to undertake the analyses in this study (Ammermueller & Pischke, 2009).

Table 2: Decomposition of variance

Mean	0.030
Between	0.110
Within	0.045
Total	0.155

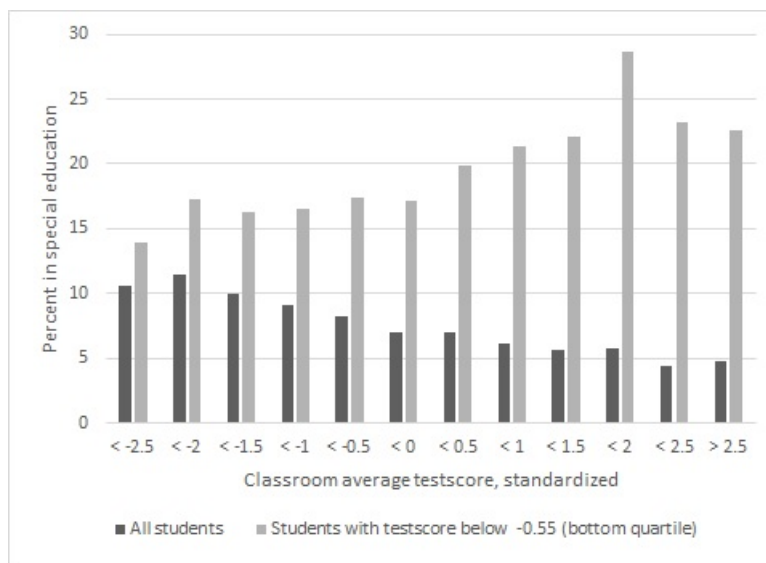


# Results

Before formally testing the hypothesis that students with better-performing classmates are more likely to receive special education regardless of their own academic performance, I will present a graphical illustration of the correlation. Figure 3 illustrates the data and the correlation of interest for this analysis in a simplified format. The x-axis is classroom average performance divided into evenly sized categories, and the Y-axis is the percentage of students in special education. The dark bars are the percentage of all students referred and the lighter bars are the percentages of students in the lowest performing quartile of the population referred to special education. The dark bars follow the pattern that one would expect if there were no inequality in service provision, namely that more students in lower-performing classes are referred to special education. The interesting part is the lighter bars, which show the opposite pattern. A higher share of students performing in the bottom quartile are referred to special education in better-performing classes than in lower-performing classes.

Thus, Figure 3 indicates that a low-performing student is more likely to receive special education if she is assigned to a classroom with better-performing students.

Figure 3: Graphical illustration of data



In line with the hypothesis, a formal test of this reveals the same pattern. The results in Table 3 show that a student is more likely to be referred to special education when she has better-performing peers. Since school-level variables are fixed and assignment to classes within schools is as good as random, this implies that the likelihood of a student being referred to special education is affected by her classmates. As can be seen in Table 3, Model 1, a student in a class that performs on average one standard deviation better in the national tests is 1.5 percentage points more likely to be referred to special education, given her own academic performance. Considering the baseline that 7.23 % of the students in the population receive special education, this effect size implies that a student in a class that performs one standard deviation worse than the average class is 20.7 % less likely to receive special education than an equally performing student in an average-performing class—a finding that is statistically significant at the 0.001 level. When background information on parental education, gender, immigrant status and the presence of psychiatric diagnoses at the individual and class level is included in Model 2, it does not change the grouping effect, which further emphasizes that this pattern is not caused by the variation between classes on observable measures.

In Model 3, an interaction term between individual performance and class-level average is included. It is positive, highly significant and about a third of the size of the class average itself. This indicates that the better a student is performing, the more class average matters with regard to the likelihood of special education referral. Or more intuitively, special education referrals for students with lower test scores are less sensitive to the classroom average. This implies that the classroom composition is most important in borderline cases. The needs of very low-performing students will often be met with special education, whereas the referral of less low-performing students will depend to a greater extent on the classroom composition.

As described above, the teacher is the gatekeeper to special education, but whether the student should receive special education, and if so, how much, is based on tests made by

Table 3: Referrals to special education

	Model 1	Model 2	Model 3
Grouping effect	0.015*** (0.001)	0.015*** (0.001)	0.015*** (0.001)
Test score	-0.076*** (0.002)	-0.073*** (0.002)	-0.073*** (0.002)
3 <sup>rd</sup> grade	0.000 .	0.000 .	0.000 .
4 <sup>th</sup> grade	-0.003 (0.002)	-0.004 (0.002)	-0.004 (0.002)
5 <sup>th</sup> grade	-0.007*** (0.002)	-0.009*** (0.002)	-0.009*** (0.002)
6 <sup>th</sup> grade	-0.011*** (0.002)	-0.013*** (0.002)	-0.013*** (0.002)
7 <sup>th</sup> grade	-0.025*** (0.003)	-0.029*** (0.003)	-0.028*** (0.003)
8 <sup>th</sup> grade	-0.032*** (0.003)	-0.036*** (0.003)	-0.035*** (0.003)
9 <sup>th</sup> grade	-0.040*** (0.003)	-0.043*** (0.003)	-0.044*** (0.003)
Parental education		-0.009*** (0.001)	-0.009*** (0.001)
Boy		0.013*** (0.001)	0.013*** (0.001)
Immigrant		-0.018*** (0.003)	-0.018*** (0.003)
Psychiatric diagnosis		0.088*** (0.004)	0.088*** (0.004)
Class share, immigrants		0.003* (0.001)	0.003* (0.001)
class share, boys		0.000 (0.001)	0.000 (0.001)
Class share, psychiatric diagnosis		0.000 (0.001)	0.000 (0.001)
Class average, parental education		-0.001 (0.001)	-0.000 (0.001)
grouping effect * testscore			0.004*** (0.001)
Constant	0.091*** (0.002)	0.083*** (0.002)	0.081*** (0.002)
N	314504	312590	312590
Schools	1330	1330	1330

Note: School fixed effects, OLS coefficients, robust standard errors in parentheses. \*p < .05; \*\*p < .01; \*\*\*p < .001.

a centrally organized PPA unit. The PPA unit covers many schools and many classes. They only meet the individual students and hence they have no knowledge of the student's classmates. Therefore, their assessments can be used to test whether the pattern revealed in Table 3 is simply because more students in better classes need special education, or whether it is a result of an influence of the administrative grouping of students.

Table 4, Model 1 shows that the amount of special education students receive is positively correlated with the average performance of the class, and that this is statistically significant at the 0.001 level. That means, when we look at all students, on average students in better classes receive more minutes of special education. However, in Model 2 the analysis is restricted to students who receive special education, and here the pattern is different. The amount of special education that a special needs student receives does not increase with the performance of the student's classmates; actually it decreases, though the coefficient is not statistically significant. Thus, the results in Table 4 imply that students in better-performing classes on average receive more special education, but that this correlation is driven by the over-representation of students from high-performing classes in special education. The special needs students in high-performing classes do not receive more special education than equally low-performing students in other classes. When students with the same academic level do not receive significantly different amounts of special education, this implies that there are not any omitted class-level variables in the analysis that drive the inequality in referrals.

This finding illustrates that when the student's needs are assessed by someone outside the class, the class becomes unimportant; but the classroom seems to matter with regard to whether a student's needs for special education will be assessed in the first place.

There is one caveat to this conclusion that needs some further consideration. Even though the comparison is between classes within the same school, and hence the compared classes draw from the same pool of resources, different classes within the school could be allocated different resources. The Danish law on special education states that special education is given

Table 4: Amount of special education

	Model 1	Model 2
	Hours of special education pr. student	Hours of special education pr. special education student
Grouping effect	0.018*** (0.002)	-0.006 (0.029)
Test score	-0.103*** (0.005)	-0.121*** (0.026)
3 <sup>rd</sup> grade	0.000 .	0.000 .
4 <sup>th</sup> grade	-0.013 (0.007)	-0.056 (0.056)
5 <sup>th</sup> grade	-0.020*** (0.006)	-0.113* (0.055)
6 <sup>th</sup> grade	-0.040*** (0.006)	-0.212*** (0.063)
7 <sup>th</sup> grade	-0.040*** (0.007)	-0.121 (0.091)
8 <sup>th</sup> grade	-0.069*** (0.007)	-0.279*** (0.079)
9 <sup>th</sup> grade	-0.081*** (0.007)	-0.382*** (0.091)
Constant	0.102*** (0.005)	1.347*** (0.045)
N	307498	17383
Schools	1330	836

Note: School fixed effects, individual and class level variables included, but not shown. OLS coefficients, robust standard errors in parentheses. \*p < .05; \*\*p < .01; \*\*\*p < .001.

to students whose "development necessitates special consideration or support that cannot be given within the setting of the ordinary class." From the analyses above, it cannot be ruled out that the *ordinary settings of class* are different between lower- and higher-performing classrooms. One could imagine that lower-performing classes are more likely to receive extra resources from the principal to support students' learning within the settings of ordinary class. With more resources to support students inside the classroom, maybe some students who would otherwise need special education can receive the necessary support in class.

As discussed above, there are two possible ways to control for resource allocation between classes with the available data. The first is to include class size as a control variable in the regression, and the second is to limit the analysis to younger students. The academic level of the students is unknown when they start, and because classes typically have the same teacher for at least three years, the teachers in lower grades are expected to be assigned more randomly to younger classes. Results of the two analyses are shown below.

Table 5: Referrals to special education, control for resources

	Model 1	Model 2	Model 3
	All students	3 <sup>rd</sup> grade	3 <sup>rd</sup> grade no siblings
Grouping effect	0.015*** (0.001)	0.021*** (0.003)	0.021*** (0.004)
Test score	-0.073*** (0.002)	-0.086*** (0.003)	-0.078*** (0.004)
Class size	-0.002* (0.001)		
Constant	0.083*** (0.002)	0.084*** (0.002)	0.079*** (0.003)
N	312590	49653	18635
Schools	1330	1283	1282

Note: School fixed effects, individual and class level variables included, but not shown. Grade indicators included in Model 1, but not shown. OLS coefficients, robust standard errors in parentheses. \* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ .

In Table 5, Model 1, class size is included in the model. The class size coefficient is negative, but quite small. In line with the argument that teachers may be better able to provide in-class support to special needs students in smaller classes than larger classes, this

result implies that students in smaller classes are slightly less likely to be referred to special education. However, the class size coefficient is small and the inclusion of class size does not alter the grouping effect, which indicates that the influence from the classroom composition is not mediated by class size. In Table 5, Model 2, the analysis is limited to 3<sup>rd</sup> grade students. Here, the grouping effect is larger than in the complete sample analysis. Thus, students in 3<sup>rd</sup> grade are more likely to receive special education in general, and their referral also seems to be influenced by the classroom to a larger extent than older students. When analysis is further limited to students who do not have older siblings and do not live in households with older children in Model 3, the grouping effect is exactly the same, which indicates that parents with knowledge of the school and the teachers do not interfere in the allocation of resources to classrooms in a way that would ensure compensation to lower-performing classrooms. These tests are not sufficient to rule out the possibility that lower-performing classes may be compensated in some way that could make referrals to special education unnecessary, but the analyses that can be made to illuminate this explanation do not find any evidence that the referral pattern in Table 3 is not caused by an influence from classroom composition.

## Concluding discussion

This study has taken the first step in theorizing about administrative grouping of citizens in street-level agencies. The empirical investigation of the influence of the administrative grouping of students in classes showed that students in better-performing classes were more likely to receive special education than equally performing students in lower-performing classes. The variation between administrative groups arose from as-good-as-random assignment of students to classes within schools. Despite small differences between classes, this study found a very large effect of administrative grouping. Thus, a student in a class performing one standard deviation lower than an average class was 20.7 % less likely to be referred to special

education than an equally performing student in an average-performing class.

In a mapping of the characteristics of students assigned to special education made by Hibel, Farkas, and Morgan (2010), it was also found that students from better-performing schools were more likely to receive special education. This mapping cannot account for selection into schools, and therefore the finding could be a result of unevenly distributed resources. In contrast, I compare classes within the same school. Since resources are allocated at school-level and all classes within the same school draw from the same pool of resources, differences in available resources to accommodate special education across classes do not explain the differential referral patterns across classes. The consistency between the mapping and this study highlights the importance of micro-level explanations.

Administrative groups may often be identifiable on different levels of an organization; in this case the class, the cohort and the entire school could be considered an administrative group. Given the micro-level arguments presented above, one would expect that group composition has the biggest impact on the lower organizational levels. However, street-level bureaucrats may have an awareness of a group on a higher or more abstract level than the low-level group identified here. While teachers teach the same class for a number of years, they often teach several classes at a time, and hence the school-level grouping may also influence their perception of how low an academic performance ought to be to necessitate special education.

Despite the limitation of the analysis to variations between classes within the school, the influence of the administrative grouping found in this analysis is quite large. At least two explanations are possible. First, there could be differences in the allocation of resources between classes within schools. This cannot be completely ruled out, but the analyses above indicate that this is not the case. Second, classes where many students are low-performing may adapt to a lower academic level in general, and thereby accommodate the needs of the low-performing students in class without referring them to special education. Despite



the reasonableness of this argument, such a procedure may have some distributive consequences that are unwarranted and currently unattended. The analysis above shows that low-performing students in better-performing classes are more likely to receive special education, which implies an allocation of resources that does not favor those who need them most.

An essential question emerges from these results, namely whether the influence of administrative grouping identified in this study is also pertinent in other public service agencies. While many organizational dimensions may affect provision of specific services, the theory section above outlined two organizational dimensions that are expected to moderate the importance of administrative grouping across agencies, namely the specialization principle and the possibility of interaction between group members.

The specialization of schools by place implies that each school will serve citizens with many different needs and demands, and also that the school must provide a variety of different services. The various tasks teachers perform in their interactions with students—referrals to special education being just one of them—may decrease the level of specialization. A higher degree of specialization may reduce the influence of the administrative group, because more specialized street-level bureaucrats have greater expertise in the specific issue at hand and are therefore expected to be less influenced by irrelevant cues like the composition of the administrative group. In this case, just about 7 % of students receive special education, but presumably more students are considered for special education by the teacher. Furthermore, a central element of a teacher's job is to assess students' progress. This assessment is a continuous and necessary task for teachers in their planning, their support of and feedback to students, their dialogue with parents and their collaboration with other teachers teaching the same class. One may therefore assume that even though teachers do not refer students to special education on a daily basis, they are somewhat specialized in identifying students who are academically lagging behind. Thus, the influence of the administrative grouping may be

larger in, but not limited to, agencies with lower levels of expertise.

Another organizational characteristic of schools is the interaction between citizens assigned to the same group. This interaction is a precondition for the demand-side mechanism. When citizens do not know their needs relative to the other group members, the group cannot inform their demands. Given other studies' implicit support for a comparative element in street-level bureaucrats' decision making, it seems unlikely that the effect above is solely a result of group-dependent demands, but future studies should try to disentangle the two mechanisms. While the interaction among citizens may change citizens' demands for public services, the interaction also entails that the street-level bureaucrats observe all group members at the same time. It seems plausible to assume that the group, and the distribution of the needs within it, becomes more salient to the street-level bureaucrat when the individual is assessed in the presence of the remaining group. However, studies of assessments made case-by-case reveal a consistent pattern of an influence from the sequence of cases (Chen, Moskowitz, & Shue, 2016; Simonsohn & Gino, 2013). Though the exact mechanism may be different in agencies where citizens interact with the street-level bureaucrats individually and not within the group, these studies indicate that other group members may influence the street-level bureaucrat's assessment even when they are not present.

There may be several ways to mitigate the distributive consequences of grouping within as well as between agencies. First, agencies can be compensated for higher needs among their clients. However, if the unfairness arises at the street level, it is important to note that equalization transfers based on local assessments of needs will not alleviate the grouping effect, but rather cause an unintended larger compensation to lower-need agencies. Therefore, for equalization transfers to have the intended effect, they need to be made on non-discretionary measures. This taps into a more general discussion of street-level bureaucrats' room for discretion. If decisions on service provision were not subject to a discretionary assessment by street-level bureaucrats, most biases held by them would become irrelevant to the de-

cisions of who gets what, when and how. This would also imply that the administrative groupings would be of no importance to service provision. However, as argued above, such increased control may come at the cost of sub-optimal policies, because the multifaceted dilemmas street-level bureaucrats find themselves in are more complex than most regulation can cover. Therefore, in most cases increased control will most likely reduce the influence of administrative grouping, but at the same time create a number of other inequalities.

Without limiting street-level bureaucrats' discretion, other measures could be taken to minimize the influence of administrative grouping. If the influence of administrative grouping is caused by street-level bureaucrats using the group for comparison, one would expect that the impact of grouping could be drastically reduced if more people were working on the same case, or if the street-level bureaucrat had experience from different administrative groupings. When the perceptions of normality, needs and eligibility are rooted in a broader frame of reference either individually or through collaboration between employees, one would expect the influence from a specific grouping to be weaker in the assessments of individual group members than if the grouping itself constituted the primary experience of normality, needs and eligibility. Means to reach a broader frame of reference could be rotation principles or mandatory work placements in organizations with different group compositions. Another dimension that could affect the influence of administrative grouping is the street-level bureaucrats' level of experience or expertise, as discussed above. Studies of sequence effects show that more experienced employees are less susceptible to the sequence of cases in their assessments (Chen, Moskowitz, & Shue, 2016), and Harrits (2019) argues that social workers operating in a more heterogeneous social context were less likely to use stereotypes in their encounters with citizens. The argument is that experience and expertise help street-level bureaucrats identify qualifications, needs and risk irrespective of the context they occur in. This implies that more specialization by function or increased education may reduce group-dependent biases.

In the case examined here, an external employee from the PPA unit assesses the needs for special education among the students recommended for special education by their teacher. This procedure does not alleviate the inequality in referrals, but within the group of students referred to special education, the external assessment of needs was not affected by the administrative grouping of students in classes. One cannot conclude from this study whether the employees in the PPA unit suffer from their own group bias arising from the group of students they assess, but the two-tiered assessment process did result in the disappearance of the group dependency found in the analysis of referrals. It cannot be concluded from this analysis whether this was due to the different administrative groups of citizens, PPA employees' greater expertise in these assessments or their broader frame of reference because they each serve multiple school districts. But the results indicate that at least in sum, collaboration, broader experience with different citizens and higher levels of expertise can reduce the influence of administrative groupings.

Inspired by our current knowledge on coping behavior, another dimension to consider in future work is the workload experienced by street-level bureaucrats. When resources are scarce, the time spent on any individual case may be shorter and the need for an easy comparative standard provided by the administrative grouping may be greater. This would imply that decreased workload could also decrease coping behavior rooted in administrative groupings.

This study contributes a new theoretical argument about differential treatment in street-level bureaucracies and an outline of different ways to mitigate this effect. Though further research is needed to make firm conclusions about when and how the administrative grouping of citizens influences the provision of public services, the presence of a simple and strong pattern like the one showed above in an expensive and complex case of service provision, with professional street-level bureaucrats and a two-tiered referral scheme, indicates that this phenomenon may exist more broadly than just in referrals of students to special educa-

tion. Increased knowledge of this phenomenon will advance our understanding of differential treatment in street-level bureaucracies and enable decision makers to ensure that the administrative grouping of citizens does not cause inequality in service provision, and may even point to ways in which the administrative grouping of citizens may be used to increase equality.

## References

- Ammermueller, A., & Pischke, J.-S. (2009). Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics*, 7(3), 315–348.
- Andersen, L. B. (2009). What determines the behaviour and performance of health professionals? Public service motivation, professional norms and/or economic incentives [Publisher: SAGE Publications Ltd]. *International Review of Administrative Sciences*, 75(1), 79–97.
- Andersen, S. C., & Guul, T. S. (2019). Reducing Minority Discrimination at the Front Line—Combined Survey and Field Experimental Evidence [Publisher: Oxford Academic]. *Journal of Public Administration Research and Theory*, 29(3), 429–444.
- Andersen, S. C., & Hjortskov, M. (2016). Cognitive Biases in Performance Evaluations. *Journal of Public Administration Research and Theory*, 26(4), 647–662.
- Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Bawn, K. (1995). Political Control Versus Expertise: Congressional Choices about Administrative Procedures [Publisher: [American Political Science Association, Cambridge University Press]]. *The American Political Science Review*, 89(1), 62–73.

- Beck, N. (2020). Estimating Grouped Data Models with a Binary-Dependent Variable and Fixed Effects via a Logit versus a Linear Probability Model: The Impact of Dropped Units. *Political Analysis*, 28(1), 139–145.
- Beuchert, L. V., & Nandrup, A. B. (2014). The Danish National Tests – A Practical Guide. *Institut for Økonomi, Aarhus Universitet. Economics Working Papers No. 2014-25*.
- Boadway, R., & Shah, A. (2009). *Fiscal Federalism: Principles and Practice of Multiorder Governance*. Cambridge University Press.
- Boone, S., Thys, S., Van Avermaet, P., & Van Houtte, M. (2018). Class composition as a frame of reference for teachers? The influence of class context on teacher recommendations [tex.publisher: Blackwell Publishing Ltd]. *British Educational Research Journal*, 44(2), 274–293.
- Brodkin, E. Z. (1997). Inside the Welfare Contract: Discretion and Accountability in State Welfare Administration [Publisher: University of Chicago Press]. *Social Service Review*, 71(1), 1–33.
- Brodkin, E. Z. (2011). Policy Work: Street-Level Organizations Under New Managerialism. *Journal of Public Administration Research & Theory*, 21(suppl\_2), i253–i277.
- Calsamiglia, C., & Loviglio, A. (2019). Grading on a Curve: When Having Good Peers Is Not Good. *Economics of Education Review*, 73.
- Chaney, C. K., & Saltzstein, G. H. (1998). Democratic Control and Bureaucratic Responsiveness: The Police and Domestic Violence [Publisher: [Midwest Political Science Association, Wiley]]. *American Journal of Political Science*, 42(3), 745–768.
- Chen, D. L., Moskowitz, T. J., & Shue, K. (2016). Decision Making Under the Gambler’s Fallacy: Evidence from Asylum Judges, Loan Officers, and Baseball Umpires [tex.publisher: Oxford University Press, UK]. *The Quarterly Journal of Economics*, 131(3), 1181.

- Davis, J. A. (1966). The Campus as a Frog Pond: An Application of the Theory of Relative Deprivation to Career Decisions of College Men [tex.publisher: University of Chicago Press]. *American Journal of Sociology*, 72(1), 17–31.
- Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, 7(2), 117–140.
- Folkeskoleloven. (2017). LBK1510 (2017). <https://www.retsinformation.dk/eli/lta/2017/1510>
- Guul, T. S., Pedersen, M. J., & Petersen, N. B. G. (2021). Creaming among Caseworkers: Effects of Client Competence and Client Motivation on Caseworkers' Willingness to Help [\_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/puar.13297>]. *Public Administration Review*, 81(1), 12–22.
- Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2002). Inferring Program Effects for Special Populations: Does Special Education Raise Achievement for Students with Disabilities? [Publisher: MIT Press]. *The Review of Economics and Statistics*, 84(4), 584–599.
- Harrits, G. S. (2019). Stereotypes in Context: How and When Do Street-Level Bureaucrats Use Class Stereotypes? *Public Administration Review*, 79(1), 93–103.
- Helson, H. (1947). Adaptation-Level as Frame of Reference for Prediction of Psychophysical Data. *The American Journal of Psychology*, 60(1), 1–29.
- Hibel, J., Farkas, G., & Morgan, P. L. (2010). Who Is Placed into Special Education? [Publisher: SAGE Publications Inc]. *Sociology of Education*, 83(4), 312–332.
- Holm, J. M. (2017). Double standards? How historical and political aspiration levels guide managerial performance information use. *Public Administration*, 95(4), 1026–1040.
- Hoxby, C. M. (2000). Peer Effects in the Classroom: Learning from Gender and Race Variation [tex.publisher: National Bureau of Economic Research]. *NBER Working paper 7867*.
- Huber, J. D., & Shipan, C. R. (2002). *Deliberate Discretion?: The Institutional Foundations of Bureaucratic Autonomy* (1st ed.). Cambridge University Press.

- Jacob, B. A. (2005). Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, 89(5), 761–796.
- Jensen, D. C., & Pedersen, L. B. (2017). The Impact of Empathy—Explaining Diversity in Street-Level Decision-Making. *Journal of Public Administration Research and Theory*, 27(3), 433–449.
- Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. *Organizational Behavior and Human Decision Processes*, 51(2), 296–312.
- Keslair, F., Maurin, E., & McNally, S. (2012). Every child matters? An evaluation of “Special Educational Needs” programmes in England. *Economics of Education Review*, 31(6), 932–948.
- Kvande, M. N., Bjørklund, O., Lydersen, S., Belsky, J., & Wichstrøm, L. (2019). Effects of special education on academic achievement and task motivation: A propensity-score and fixed-effects approach. *European Journal of Special Needs Education*, 34(4), 409–423.
- Lavy, V., & Sand, E. (2018). On the origins of gender gaps in human capital: Short- and long-term consequences of teachers’ biases. *Journal of Public Economics*, 167, 263–279.
- Lipsky, M. (1980). *Street-level bureaucracy : Dilemmas of the individual in public services* (30th anniversary expanded edition: 2010.) [ISSN: 9780871545442]. Russell Sage Foundation.
- Marsh, H. W. (1984). Self-concept: The Application of a Frame of Reference Model to Explain Paradoxical Results [tex.publisher: SAGE Publications Ltd]. *Australian Journal of Education*, 28(2), 165–181.
- Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O’Mara, A. J., & Craven, R. G. (2008). The Big-Fish-Little-Pond-Effect Stands Up to Critical Scrutiny: Im-



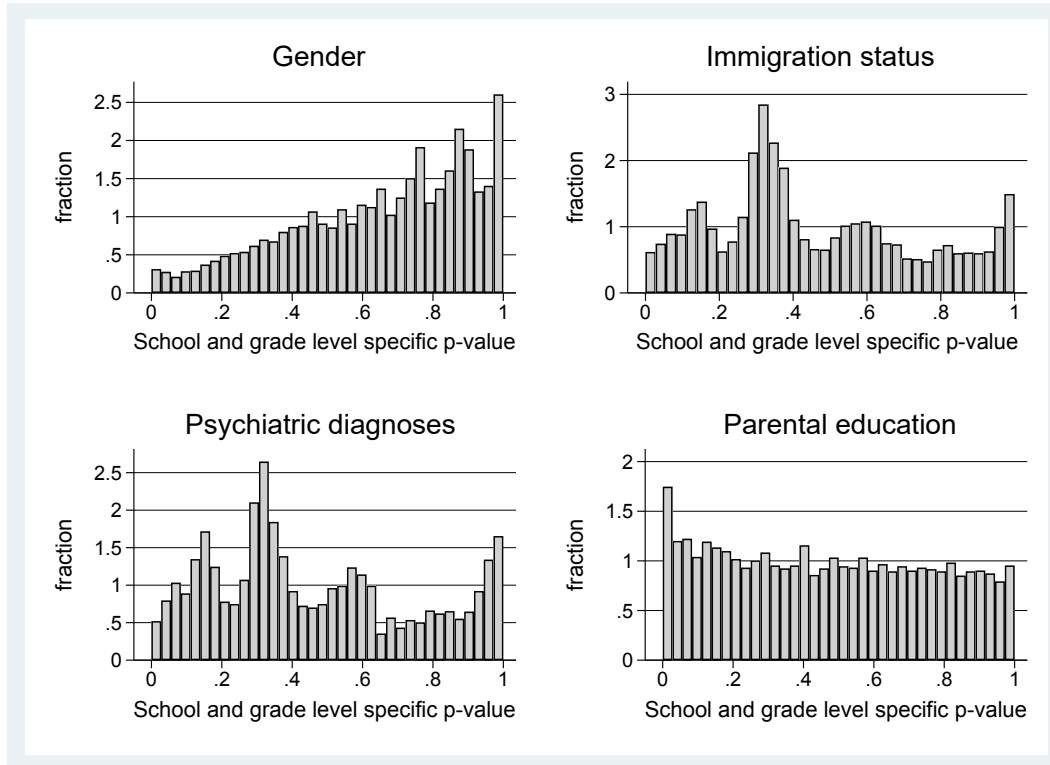
- plications for Theory, Methodology, and Future Research [tex.publisher: Springer]. *Educational Psychology Review*, 20(3), 319–350.
- Maynard-Moody, S., & Musheno, M. C. (2003). *Cops, teachers, counselors: Stories from the front lines of public service* [tex.lccn: HV40.8.U6 M39 2003]. University of Michigan Press.
- Meier, K. J., & O’Toole, L. J. (2006). Political Control versus Bureaucratic Values: Reframing the Debate [Publisher: [American Society for Public Administration, Wiley]]. *Public Administration Review*, 66(2), 177–192.
- Ministry of Finance. (2010). *Specialundervisning i folkeskolen - veje til en bedre organisering og styring* (tech. rep.). Copenhagen.
- Møen, J., & Tjelta, M. (2010). Grading Standards, Student Ability and Errors in College Admission. *Scandinavian Journal of Educational Research*, 54(3), 221–237.
- Møller, A. M. (2021). Deliberation and Deliberative Organizational Routines in Frontline Decision-Making. *Journal of Public Administration Research and Theory*, 31(3), 471–488.
- Neumann, M., Trautwein, U., & Nagy, G. (2011). Do central examinations lead to greater grading comparability? A study of frame-of-reference effects on the University entrance qualification in Germany. *Studies in Educational Evaluation*, 37(4), 206–217.
- Nielsen, P. A. (2014). Learning from Performance Feedback: Performance Information, Aspiration Levels, and Managerial Priorities. *Public Administration*, 92(1), 142–160.
- Olsen, A. L. (2017). Compared to What? How Social and Historical Reference Points Affect Citizens’ Performance Evaluations. *Journal of Public Administration Research and Theory*, 27(4), 562–580.
- Olson, J. C. (2016). Race and Punishment in American Prisons. *Journal of Public Administration Research and Theory*, 26(4), 758–768.

- Pedersen, M. J., Stritch, J. M., & Thuesen, F. (2018). Punishment on the Frontlines of Public Service Delivery: Client Ethnicity and Caseworker Sanctioning Decisions in a Scandinavian Welfare State. *Journal of Public Administration Research and Theory*, 28(3), 339–354.
- Sandfort, J. R. (2000). Moving Beyond Discretion and Outcomes: Examining Public Management from the Front Lines of the Welfare System. *Journal of Public Administration Research and Theory*, 10(4), 729–756.
- Seaton, M., Marsh, H. W., & Craven, R. G. (2010). Big-Fish-Little-Pond Effect: Generalizability and Moderation - Two Sides of the Same Coin [tex.publisher: [American Educational Research Association, Sage Publications, Inc.]]. *American Educational Research Journal*, 47(2), 390–433.
- Simon, H. A. (1939). The administrator in search of statistics. *Public Management*, 21, 106–109.
- Simon, H. A. (1945). *Administrative Behavior* (4th). The Free Press.
- Simonsohn, U., & Gino, F. (2013). Daily Horizons: Evidence of Narrow Bracketing in Judgment From 10 Years of M.B.A. Admissions Interviews. *Psychological Science*, 24(2), 219–224.
- Soss, J., Fording, R. C., & Schram, S. (2011). The Organization of Discipline: From Performance Management to Perversity and Punishment. *Journal of Public Administration Research and Theory*, 21, i203–i232.
- Specialundervisningsloven. (2010). LBK 885 <https://www.retsinformation.dk/eli/lta/2010/885>
- Thomann, E., van Engen, N., & Tummers, L. (2018). The Necessity of Discretion: A Behavioral Evaluation of Bottom-Up Implementation Theory. *Journal of Public Administration Research and Theory*, 28(4), 583–601.
- Tiebout, C. M. (1956). A Pure Theory of Local Expenditures [Publisher: The University of Chicago Press]. *Journal of Political Economy*, 64(5), 416–424.

- Tummers, L. (2017). The Relationship Between Coping and Job Performance. *Journal of Public Administration Research and Theory*, 27(1), 13.
- Tummers, L., Bekkers, V., Vink, E., & Musheno, M. (2015). Coping During Public Service Delivery: A Conceptualization and Systematic Review of the Literature [Publisher: Oxford Academic]. *Journal of Public Administration Research and Theory*, 25(4), 1099–1126.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131.
- Weber, M. (1922). *Economy and Society*. Bedminster Press Incorporated: 1968.
- Weber, M. (1947). *The theory of social and economic organizations*. Oxford University Press.
- Wenger, J. B., & Wilkins, V. M. (2008). At the Discretion of Rogue Agents: How Automation Improves Women’s Outcomes in Unemployment Insurance. *Journal of Public Administration Research and Theory*, 19(2), 313–333.

# Appendix

Figure A1: Distribution of school cohort-specific p-values for Pearson Chi<sup>2</sup>



Evenly distributed p-values indicates random assignment. A right-skewed distribution indicates less difference than would be expected from random allocation, whereas a left-skewed distribution indicates bigger differences than a completely random allocation would create (Ammermueller & Pischke, 2009)

Table A1: Referrals to special education, complete cohorts only

	Referral
Grouping effect	0.020*** (0.002)
Test score	-0.081*** (0.003)
3 <sup>rd</sup> grade	0.000 .
4 <sup>th</sup> grade	-0.001 (0.006)
5 <sup>th</sup> grade	-0.009* (0.005)
6 <sup>th</sup> grade	-0.018** (0.007)
7 <sup>th</sup> grade	-0.032*** (0.006)
8 <sup>th</sup> grade	-0.040*** (0.008)
9 <sup>th</sup> grade	-0.048*** (0.007)
Constant	0.087*** (0.003)
N	66589
Schools	888

Note: School fixed effects, individual- and class-level variables included, but not shown. OLS coefficients, robust standard errors in parentheses. \*p < .05; \*\*p < .01; \*\*\*p < .001.

Table A2: Referrals to special education, conditional logistic regression

	Referral
Grouping effect	0.277*** (0.012)
Test score	-1.251*** (0.010)
3 <sup>rd</sup> grade	0.000 .
4 <sup>th</sup> grade	-0.086*** (0.028)
5 <sup>th</sup> grade	-0.153*** (0.026)
6 <sup>th</sup> grade	-0.237*** (0.028)
7 <sup>th</sup> grade	-0.525*** (0.028)
8 <sup>th</sup> grade	-0.660*** (0.031)
9 <sup>th</sup> grade	-0.895*** (0.033)
N	254397
Schools	1044

Note: School fixed effects, individual- and class-level variables included, but not shown. Odds ratios, standard errors in parentheses. \*p < .05; \*\*p < .01; \*\*\*p < .001.

Table A3: Referrals to special education, Linear probability model estimated with conditional logistic regression sample

	Referral
Grouping effect	0.018*** (0.001)
Test score	-0.091*** (0.002)
3 <sup>rd</sup> grade	0.000 .
4 <sup>th</sup> grade	-0.005 (0.003)
5 <sup>th</sup> grade	-0.011*** (0.003)
6 <sup>th</sup> grade	-0.017*** (0.003)
7 <sup>th</sup> grade	-0.035*** (0.003)
8 <sup>th</sup> grade	-0.044*** (0.003)
9 <sup>th</sup> grade	-0.054*** (0.003)
Constant	0.102*** (0.002)
N	254397
Schools	1044

Note: School fixed effects, individual- and class-level variables included, but not shown. OLS coefficients, standard errors in parentheses. \*p < .05; \*\*p < .01; \*\*\*p < .001.