Journal of Public Economics 188 (2020) 104204

Contents lists available at ScienceDirect

Journal of Public Economics

journal homepage: www.elsevier.com/locate/jpube

Child care center quality and early child development☆

Nina Drange [a,b,*], Marte Rønning [b]

[a] Ragnar Frisch Centre for Economic Research, Norway
[b] Statistics Norway, Norway

ARTICLE INFO

Article history:
Received 17 June 2019
Received in revised form 1 May 2020
Accepted 5 May 2020
Available online 21 May 2020

JEL classifications:
I21
J13

Keywords:
Child care quality
Child development

ABSTRACT

A unique dataset on applications and admissions to child care centers allow us to explore parental preferences for child care center characteristics. We proceed to study how staff qualifications, experience and sickness absence, as well as the proportion of male and immigrant staff, explain the cognitive development of children randomly allocated to child care centers contingent on observable characteristics. Children who receive their first offer of child care enrollment from a center with a high share of male staff, perform better on language tests in their early school years. If the sickness absence is high in the center, child test scores are lower in both language and mathematics.

1. Introduction

It is well-documented that intensive and high-quality child care improves the lives of deprived children (Almond and Currie, 2011; Baker, 2011; Ruhm and Waldfogel, 2012; Blau and Currie, 2006). However, it is still unclear which components of a child care program enhance child development (Blau and Currie, 2006). While studies from the U.S. seem to largely dismiss the role of structural quality, such as the share of educated teachers and group size (Blau, 1999; Currie and Neidell, 2007; Walters, 2015), studies from European countries suggest that structural indicators, as well as the teacher gender composition, may improve child outcomes (Bauchmuller et al., 2014; Goertz et al., 2018). Moreover, studies of process quality, reflecting the interaction between the child and its caregivers (Blau and Currie, 2006), indicate that teachers vary substantially in effectiveness (Araujo et al., 2016). Given the surge in enrollment of children in child care[1] across most OECD countries during the last decade, as well as the quantity of subsidies paid by many governments, it appears to be of great importance that we improve our understanding of how child care quality can advance child outcomes.

Taking advantage of child care assignment lotteries, a recent study of child care in Norway reports that enrolling children aged 1–2 in child care has positive impacts on their first-grade performance in language and mathematics (Drange and Havnes, 2019). We rely on a similar identification and data to study how parents value structural quality inputs in child care centers, and how these inputs explain child test scores during early school years. Failure to take into account the parental selection of centers will bias results, for example if parents who provide a home environment fostering positive child development select child care centers that do so as well. We address this challenge by exploiting a unique data set from the municipality of Oslo containing applications to child care centers, in addition to a system of quasi-random assignment. During the years covered by our data, child care centers in Oslo were oversubscribed, and child care slots were allocated through a lottery. While most children who wanted to attend a child care center would eventually enroll, the majority ended up enrolling in a center other than the preferred one. This allows us to compare the development of children whose parents initially applied for the same center(s), but where the children ended up being assigned to centers with differing structural quality as a result of oversubscription. Our data allow us to assess the role of a wide range of structural child care center inputs, such as the characteristics of child care center employees, as well as group and center size. Moreover, we have information on the distance between the child's home and the center. We measure the cognitive development of a child at age 6–9, drawing on a pooled measure of test scores in the first, second and third grades.

Our paper adds to the literature along three important dimensions. First, the combination of rich data and the institutional setting allow us to plausibly control for parental preferences and application

☆ Thanks to Oslo Municipality for generously providing data, institutional detail and feedback. We also wish to thank Kjetil Telle, Edwin Leuven, Ingrid Huitfeldt and two anonymous referees for comments and suggestions to earlier versions of the paper. The project received financial support from the Norwegian Research Council (Grant Number 236947).
* Corresponding author at: Ragnar Frisch Centre for Economic Research, Norway.
E-mail addresses: n.e.drange@frisch.uio.no (N. Drange), mro@ssb.no (M. Rønning).
[1] Throughout the paper, when we refer to child care, we mean formal center-based care prior to school start.

https://doi.org/10.1016/j.jpubeco.2020.104204

behavior. This differs from the majority of existing papers, which often rely on within child care variation.[2] An exception is Walters (2015), who exploits the random variation from the implementation of the Head Start Impact Study and finds that centers offering full-day service and home visiting are more effective, while curriculum, teacher education and class size do not predict effectiveness. We study a different set of structural inputs, and our panel data allows us estimate models that account for center fixed-effects.

Second, we draw on rich registry data that enables us to study how parents value a wide range of structural quality measures, such as the center and group size, staff qualifications, experience and stability, as well as the proportion of male and immigrant staff. We examine how these center-related attributes relate to child test scores during the early school years, taking into account how parents select child care institutions.

Finally, we provide evidence from an understudied part of the education system. Studies of quality in schools flourish (Krueger, 1999; Clotfelter et al., 2009; Dee, 2005, 2007; Antecol et al., 2015; Rivkin et al., 2005; Hoffmann and Oreopoulos, 2009), but it is unclear whether the results can be applied to child care centers. There has long been a consensus among psychologists, neurobiologists and economists that investing in early childhood is imperative, as this is a particularly sensitive period when the child's brain is at its most receptive, and the foundation for cognitive and socio-emotional capacities is developed (Knudsen et al., 2006). Thus, enhancing our understanding of how the center environment relates to a child's cognitive development at this early stage of the education process appears to be of great importance.

Our findings suggest that parents prefer child care centers that are situated close to home, and that the likelihood of preferring a center with a higher share of pre-school teachers increases with father's income. Turning to the children, we show that those who receive an offer of enrollment in a child care center with a higher share of male staff perform better on both language and mathematics tests in their early school years. Low sickness absence among the staff also predicts positive child development. Note that while these findings are not biased by parental selection, observable inputs may correlate with unobservables in a center. We go on to estimate center fixed-effects, keeping constant unobservable center characteristics such as localities or a specific pedagogical approach. This approach yields similar estimates, supporting a causal interpretation of the result.

The paper proceeds as follows. We first present the institutional background in Section 2, before describing our data in Section 3. In Section 4 we present and discuss our empirical approach. Our main results are presented in Section 5, while Section 6 provides a conclusion.

2. Institutional setting

2.1. Child care in Norway

Child care in Norway is heavily regulated, with provisions on staff qualifications, number of children per employee and per teacher, size of play area, and educational orientation. Institutions are run by an educated child care teacher responsible for day-to-day management and educational content. The child care teacher education is a three-year college degree that includes supervised practice in a child care institution. Child care regulations specify that there should be at least one educated child care teacher per 7–9 children aged less that three years, and at least one per 14–18 children aged 3–6. Municipal regulations specify that there should be at least one employee per three children under the age of three, and one employee per six children aged 3–6. There are no educational requirements for the additional staff. Given that the child care center meets the regulations the teacher/employees ratio will this be 1 teacher per 3 employees. Few public child care centers in Norway accept children who are younger than a year old.

Enrollment rates are high in Norway. According to numbers from the OECD (2017),[3] about 55% of children under the age of 3 and 97% of children aged 3–5 were enrolled in a child care center. This was similar to Sweden and Denmark, but higher than the OECD average of 34% for under 3 year-olds and 85 for 3–5 year-olds. In Germany, 37% of the under 3 s and 97 of the over 3 s were enrolled, while the corresponding numbers for the US were 28 and 67%, respectively. When it comes to spending on child care (per child), the figure for Norway amounts to about USD 14,700. This is high compared to most countries: Germany spends USD 9200, the US about USD 10,000 and the OECD average is USD 7900. However, the high spending likely also reflects the high share of young children enrolled in Norway, as young children typically demand more resources (such as a lower child-to-staff ratio). Spending in Denmark, with the highest enrollment rate of young children in the OECD at 65%, is even higher at about USD 16,300. The child-to-teacher ratio in Norway is similar to that in Denmark, Germany and the US at 10, but lower than the OECD average of 14.

In Oslo, about 60% of child care institutions are public, whereas the remaining centers are privately operated. Both public and private institutions require municipal approval and supervision in order to be entitled to federal subsidies that cover around 80% of costs. Moreover, each enrolled child with a minority background triggers an additional subsidy to accommodate language learning. Parental payment has been capped since 2003 at around NOK 2400 per month. Child care institutions are typically open from around 7.30 am to 5 pm.

In terms of educational content, a social pedagogy tradition has dominated child care practices in Norway since the 1970s. According to this school of thought, children should develop social, language and physical skills mainly through play and informal learning.[4] The informal learning typically takes place in the context of day-to-day social interaction between children and staff, in addition to specific activities for different age groups.

In 2006, The Ministry of Education issued an overall plan for the Norwegian child care centers that is of relevance for the years our data covers (Norwegian Ministry of Education, 2006). This plan covered six focal areas, as well as a list of more specific themes and related activities that should be implemented in child care centers. While this plan is not particularly concrete in its description of learning goals and age-specific activities, it still gives an overview of what child care centers should emphasize in their pedagogical and practical work with the children. The focal areas were listed as 1) care and nurture, 2) play, 3) learning, 4) social competence, 5) language competence and 6) the child care center as a cultural arena. As for the more specific themes and activities, there is a clear emphasis on learning through play and through interaction with other children and the adults in the center. There is also a special emphasis on physical activity and the development of motor skills through both indoor and outdoor play.

2.2. Child care quality

Child care quality can be measured along a number of dimensions, but below we present certain structural aspects that theory and/or empirical evidence have shown to be of importance, and that we can explore with our data.

[2] Bauchmuller et al. (2014) estimate correlations between center inputs and child outcomes, Blau (1999) relies on mother fixed-effects models and Goertz et al. (2018) depend on within-center differences across time.

[3] Based on statistics for 2013, 2014 and 2015.

[4] The social pedagogy tradition for early education has been especially influential in the Nordic countries and Central-Europe. In contrast, a so-called pre-primary pedagogic approach to early education has dominated many English and French-speaking countries, favoring formal learning processes designed to meet explicit standards for what children should know and be able to do before they start school.

2.2.1. Group size

A lower child-to-staff ratio means that the individual child has more time with adults working in the center. As explained above, this is regulated by the municipality, and is meant to ensure that there are a sufficient number of adults in a center to care for the children in a safe setting. While it seems self-evident that a certain number of adults is required to care for young children, it is less obvious what the specific ratio should be. A meta-analysis finds that small-group instruction correlates with positive child outcomes (Camilli et al., 2010), and a similar correlation is found in a study from Denmark (Bauchmuller et al., 2014). To learn about causal effects, we need to consider literature from the school setting. Project STAR randomly assigned kindergarten students and their teachers to classes of differing size. Krueger (1999) finds that performance on standardized tests increase during the first year students attend small classes.

2.2.2. Child care teachers

The teacher in a group is responsible for the day-to-day educational approach, and the regulation in place by the time of our study was intended to ensure that each group of children had at least one teacher (and bigger groups had more, see Section 2.1). Bauchmuller et al. (2014) find a positive correlation between a higher share of teachers and child outcomes for Denmark, whereas evidence from the US largely dismisses the importance of formal qualifications in the child care setting (Blau, 1999; Currie and Neidell, 2007; Walters, 2015). For school children, Rivkin et al. (2005) study the variance in teacher quality based on within-school heterogeneity, and find that a substantial amount of the variation in child test scores is attributable to the teacher, but concludes that observable teacher characteristics such as education or experience do not predict better child outcomes.

2.2.3. Tenure in a center

Tenure is a measure that may proxy at least two different quality aspects in a center. First, longer tenure in a center presumably gives a caregiver a more child-specific knowledge, and this may be positive for child-caregiver relationships (Horm et al., 2018; Choi et al., 2019) (see also sick leave below). Second, longer tenure could be a sign of a healthy work environment, suggesting that the employees are happy working in the center. If a more content staff influences the learning environment, this in turn could be positive for child development.

2.2.4. Male share

Norwegian child care centers are mainly staffed by women, and Oslo is no exception with about one in ten staff members being male. Dee (2007) summarizes that male teachers can influence children's engagement or behavior by acting as role models, and that same-gender teachers can be positive for child development if men communicate different expectations to boys and girls. Empirical evidence from school settings does suggest that teacher gender can explain differing school performance for boys and girls (Dee, 2005, 2007; Antecol et al., 2015). Dee (2007) shows that while both girls and boys benefit from having a male mathematics teacher, girls benefit and boys perform worse if the English teacher is female, and concludes that changing an English teacher from female to male would reduce the gender gap substantially among 13 year-olds. For primary school children, Antecol et al. (2015) take advantage of data on random assignment of teachers across classrooms and schools. They find that girls assigned to female teachers suffer from lower test scores in math by the end of the academic year. Boys' results are not affected by the gender of the teacher.

Another channel through which male staff may potentially influence child outcomes in the child care center is in the play situation. The development psychology literature hypothesizes that certain types of physical play, such as rough-and-tumble play,[5] facilitates social skills practice and aggression regulation (Storli and Sandseter, 2017). This type of play is seen more often among boys, and is initiated to a greater extent by fathers than by mothers (Fletcher et al., 2013; Pellegrini and Smith, 1998). Storli and Sandseter (2017) find that both male and female child care staff promote such play in Norwegian child care centers, but male staff are more likely to have first hand experience with it. Moreover, female staff express in interviews that they have learned to facilitate rough-and-tumble play from male colleagues (Storli and Sandseter, 2017). As noted by Goertz et al. (2018), even if male staff do interact more with boys and facilitate more boy-oriented play, it is not clear whether this will affect boys more than girls. If the presence of male staff in the play situation encourages aggression regulation among boys, both girls and boys will presumably benefit from a less disruptive learning environment.

2.2.5. Immigrant share

It has been hypothesized that teachers may be more responsive to the needs of children who share their racial or ethnic background (Dee, 2004). Dee (2004) finds for the US that both black and white children benefit from a same-race teacher, while results for Denmark in Bauchmuller et al. (2014) suggest that ethnic minority children gain significantly less from a higher share of ethnic minority staff than children without such background.

2.2.6. Sick leave

Norway has a high absence due to sick leave, and child care personnel are among the employees with the highest sickness absence. A number of factors may contribute to the high sick leave among child care employees. Child care staff work closely with young children and are thus exposed to viruses on a daily basis. For employees working with young children, there is a lot of heavy lifting and sitting in uncomfortable positions on the floor.[6] Moreover, sick leave in Norway is particularly high among women, and about 90% of child care personnel are female. In the psychology literature there is a large strand of research, both theoretical but also increasingly empirical, emphasizing the importance of stable child-caregiver relationships during early childhood for later development (Bowlby, 1969; Sroufe et al., 2010). In recent years, a growing literature has studied how the relationships between young children and their non-parental care providers relate to children's later behavioral and socio-emotional functioning (Ahnert et al., 2006). Importantly, continuity of care provider is found to be related to fewer behavioral problems and higher social competence in children (Horm et al., 2018; Choi et al., 2019). We study a setting where children enter child care at an early age. If caregivers in the center have a high sickness absence, this may hamper the child's ability to form a safe attachment to its caregivers. In addition to the attachment aspect, if there are fewer staff at work than originally planned, there might be less time to provide a stimulating environment for the children. Additional activities such as more organized play or outdoor expeditions will likely be less of a priority, as will one-on-one interactions with the children. Clotfelter et al. (2009) finds that teacher absences in primary school are clearly associated with lower student achievement.

2.3. Applications and admissions to child care in Oslo

Oslo is divided into 15 city districts with their own local administrations. During the years our data covers, child care slots were allocated within the child's city district of residence. Available slots were allocated to children from other city districts if there was undersubscription in that particular district. The Oslo municipality administration handled

[5] Rough-and-tumble play is characterized by vigorous behavior (such as wrestling or play fighting) that appears to be aggressive except for the playful context (Pellegrini and Smith, 1998).

[6] According to self-reported survey data, the three most common health issues experienced by child care teachers are back pain (50% report having experienced this the previous month), neck pain (49%) and headache (44%), whereas mental health issues are reported less often (17%) (National Institute of Occupational Health 2016).

the allocation of child care slots in collaboration with the city district administrations.

A majority of the slots in the child care centers become available in August due to the transition of the six year-olds from child care to school. As a first step to obtaining a slot in August in a particular year (in both public and private centers), parents need to apply before March 1st in the same calendar year. During the time period we consider (2005–2010), parents could rank a maximum of seven child care centers.[7] Children of single mothers, disabled children and occasionally children with immigrant background were awarded priority. Subsequently, the allocation of slots in public institutions (to children with no priority) was decided in a computer-generated lottery. According to representatives from the municipality, this lottery ensured that each child care center with available slots was matched randomly with children of the appropriate age who had ranked it as one of their seven prioritized centers.[8] The main allocation process consisted of three sequential rounds, where slots that were not accepted by families with an offer of enrollment in the first round, were allocated to new children in the second round, and then similar in the third round. Drange and Havnes (2019) show that background characteristics are balanced across samples of lottery winners and losers, suggesting that the randomization was successful.

Once a family received an offer of enrollment, the child was taken out of the public lottery. However, a family could uphold their application to the first ranked center if they were offered a slot in a lower ranked center. If there were available slots in centers after all children who had applied (for those centers) had been given a slot, the municipality would offer these to children who had not yet received a slot. Towards the end of the period the lottery changed somewhat, and from 2009 the child's birth date decided the lottery number. The priorities remained the same. Contingent on these observables, the allocation of slots remained random.

In this set-up, the children could get the first offer of a slot in any of their up to seven ranked child care centers with a similar probability independently of the ranking, given similar oversubscription rates. If the family chose to remain on the waiting list for the first ranked center, they could receive a new offer in this center in a later allocation round. As we will elaborate on in Sections 3.1 and 4, we always focus on characteristics of the first offered center, regardless of which round the parents received the offer. Children not admitted at all in the main allocation round were put on a waiting list, and would only get an offer if already admitted children declined the slot they were offered.

While private and public child care centers had the same application deadline and children could apply to a mix of public and private centers, their intake rules differed somewhat. Every private child care center in Oslo received lists with detailed information on all children who had applied for a slot in that particular (private) center, including their respective ranking of the center. Subsequently the respective private institutions handled their own admissions.

Due to substantial oversubscription during this period, the majority of children received an offer from a center other than their first choice. This is documented in Table 1. In our sample, about 29% got an offer from their first ranked center, whereas about 13% got an offer from the second ranked. As many as 31% of the children in our sample were offered a center outside their choice set. The difference between being offered the first and second ranked centers is substantial, and may look puzzling at first.

Table 1
Percentage of children who get an offer from their n'th choice.

Choice	Percent
None	30.85
1st choice center	28.69
2nd choice center	13.47
3rd choice center	9.43
4th choice center	5.52
5th–7th choice center	12.05

N = 2175 children.

There are two features of the allocation mechanism that can contribute to this pattern. 1) As described above, a family who received an offer from a lower ranked center, could uphold their application to their center on first rank throughout the three lottery rounds. Hence, if the family was indifferent between being admitted to for instance three centers, but preferred these centers over other centers on their list, it would make sense to rank the center with expected lowest oversubscription first to make sure they had two draws in a center with a high(er) likelihood of being admitted. Although oversubscription was not directly observable, the information about whether there were many school starters in a center (and hence more open slots in August) was in most cases available for a parent who visited a center prior to applying. Constructing a measure for oversubscription and comparing parents that did and did not get an offer from their first choice center reveals that, perhaps unsurprisingly, parents who received an offer from their highest ranked child care center had indeed listed a center with a lower oversubscription as their first choice.[9] 2) Parents that ranked fewer centers, were more likely to secure a slot in one of the centers they had ranked, but also had a higher likelihood of being offered a center outside their choice set.[10] As explained above, children had the same probability of obtaining a slot in whatever center they had listed (given similar oversubscription rates), but the moment they were allocated a slot, they were out of the lottery.[11] A family that only ranked one or two centers would therefore have a lower probability of being out of the lottery when their first choice center entered the draw, as they had not yet been assigned to a center. In our sample, 24% listed 3 or fewer centers.

A general concern with allocation mechanisms is that they may spur strategic application behavior. In the case of the allocation mechanism we are considering, we need to take into account which institutions the families applied for, and how many centers they listed. Including these controls allows us to compare children with similar applications. This is dicussed further in Section 4.

3. Data

3.1. Dataset and variables

To conduct the analysis we employ data from several sources that can be linked through a personal identifier. Firstly, we have access to a unique data set from the municipality of Oslo containing individual records of all institutional child care use for children born between 2004 and 2007, as well as test scores from 1st to 3rd grade for the same cohorts.[12] This data set also includes full information on application

[7] Today, parents can rank five centers.

[8] The information about the lottery is based on online information about the public admission procedure (see https://lovdata.no/dokument/LTII/forskrift/2002-12-18-1831) as well as a meeting with representatives from the municipality that handled the admissions (a summary from this meeting is available upon request). We do not have access to the exact algorithm the computer was running. It should be noted, however, that the lottery results were regarded as public information and had to be given to parents who requested them. Moreover, the allocation of child care slots was a popular topic for the local newspapers (see for instance *Aftenposten Aften* 27.04.2005). Thus, the transparency should ensure that public slots were indeed allocated through the lottery mechanism.

[9] The measure of oversubscription is based on all applicants who have listed a center on any rank, by year, and a variable with the number of children admitted to a center by year. Oversubscription is defined as the former divided by the latter. These families also in general list less oversubscribed centers (than families that did not get their first ranked center), but the highest ranked center has markedly less oversubscription. Results are available from the authors upon request.

[10] Results are available from the authors upon request.

[11] Although they could still be on the waiting list for their first ranked center, and by the end of the allocation round they could have been admitted to this center. However, as mentioned above, we focus on the very first offer to each child in this analysis.

[12] Due to a restrictive storage policy in Oslo municipality, data on children born in January and February 2004 were deleted from the application data base before we gained access to it. We are therefore not able to include these children in our sample.

dates and parental preferences for child care centers. All applications ever submitted for a child are registered in the data, as is every offer of a slot the child receives, as well as identifiers for the up to seven centers the parents may rank in each application.

Children can attend a child care center in another city district than the one they reside in, but as long as the center is situated in Oslo the enrollment will be included in our data. If the child attends a child care center in another city we will not be able to register the enrollment, but this will only involve a few children. As described in Section 2.3, private child care centers have their own admission process. Children with a private institution ranked first on their application are therefore excluded from our analysis, following Drange and Havnes (2019). We also exclude children with priority as they do not participate in the lottery. In order to avoid the complication that experience with previous child care centers affects parent's ranking of centers, we focus on the first time the parents apply and the first center from which the child receives an offer. This leaves us with a sample consisting of 2175 children enrolled in 360 child care centers from 2005 to 2010.[13]

We sample child care characteristics from the first center that was offered to a family. The staff working in the different child care centers can be identified from Statistics Norway's employer-employee register (AAreg). This register includes information on the staff's experience/tenure and workload, and is merged with data on sick leave from the Norwegian Labour and Welfare Administration (NAV). These data record employees and their sick leave as of October every year. From the population, income and education registries, we collect information on staff characteristics such as gender, birthday and education. Moreover, the child care centers receiving public subsidies (almost all existing centers) must report key statistics to Statistics Norway every year, such as the number of children enrolled by age and several staff background measures. We collect the number and age of children enrolled from this registry. Since we know which center the child received its first offer from, we can link this information to each child.[14] To construct quality measures at the child care center level, we average staff characteristics (for each year) across institutions and weight the results with the staff workload. We focus on the following variables: child/employee ratio, the share of staff with a child care teacher degree, average years the current employees have been working in a given center (referred to as tenure), share of male child care staff, share of staff with immigrant background (share of individuals with both parents born abroad), share of days the staff has been absent due to long term sick leave (spells lasting more than 10 days)[15] and the size of the center measured by number of children enrolled. Since employee-data are measured in October, we base our child care quality measures on characteristics of the institution the child receives an offer from in the spring of the same year, and in most cases enrolls in in August (i.e. about two months prior to the characteristics being measured).

Fig. 1 provides an overview of the distribution of center characteristics. There is considerable variation across all characteristics, but common to most of the variables is that the distribution is skewed to the left. In the upper right figure, we report the distribution of the share with a child care teacher degree. Most centers have about one child care teacher per three employees, but many have fewer, and the average is one in five. A large share of centers do not have any men on their staffs (upper right figure). Few centers have more than 20% male staff. In the majority of centers, average employee tenure is less than five years, and the average is three (center left figure). The share of long term sick leave during the year is displayed in the center right figure. In the bottom left figure we note that the vast majority of centers have some staff with immigrant background, but that there is considerable variation. On average, the share of staff with immigrant background is 25%. Lastly, in the bottom right figure, we display child care center size (measured by the number of children enrolled in a center). We see that there are few very big centers, and few very small ones, but still a considerable variation in size.

Our preferred specification includes continuous measures of center inputs. When we explore the robustness of our results, however, we want to account for the nature of the skewed distribution. Therefore, in addition to averaging staff characteristics within centers, we also generate dummy variables for whether specific characteristics of the child care center exceed a certain threshold (the 50th and 75th percentiles in the population of child care centers).

Information on the background characteristics of the children and their families is obtained from registers provided by Statistics Norway. The covariates are measured for the year before the child was born to ensure that they are not endogenous to the treatment. Important control variables are parents' education, (net) income of the father, immigrant background and mother's continent of origin. We also control for whether parents were young (defined as being under the age of 22) when they had their first child. Furthermore, we include dummies for the child's gender, and information about birth cohort and month of birth. Lastly, we control for application year, city district and whether the child's first child care center offer is public or private. Summary statistics and further definitions of the background characteristics of the children are reported in Appendix Table 10.[16]

Norwegian children have nationwide tests in language and mathematics during their first, second and third years in school. We employ the results of these tests to construct child cognitive outcomes, and Appendix Fig. 2 shows the distribution of the outcome variables. The tests are intended to identify the weakest pupils, to ensure that the school allocates resources to underperforming children. Hence most children score close to maximum points. Due to the skewed distribution of these tests, we generate four outcome variables for each subject. The first is simply the pooled (grades 1 to 3) average achievement level in language and mathematics, while the remainder are dummy variables which are equal to one if the child scores above the 25th, 50th and 75th percentiles. In the analysis we standardize the achievement scores within cohorts and grades.

3.2. Application behavior

In order to obtain a better picture of what parents value in a child care center, we report the mean of the child care quality inputs listed by parents' rankings in Table 2. In addition to the explanatory variables listed in Appendix Table 10, Table 2 also includes average distance from

[13] Applications from 2008 are excluded from the sample, as the allocation mechanism that was in place that year could spur strategic behavior.

[14] About 30% of the child care centers in Oslo are so-called family child care centers, and few of these centers report to the employer-employee register. These centers are typically consisting of up to five children, are located in a private home and are run by an assistant under the weekly supervision of a child care teacher. Since we in most cases do not have information about their teachers, children attending these child care centers are excluded from the analysis (about 10% of the children).

[15] For each center we calculate our measure of sick leave by dividing total number of days per year the staff has been absent due to long term sick leave by total number of contractual working days.

[16] To get a better picture of whether children in our sample differ substantially from the average child care applicant, we present summary statistics for samples subject to varying restrictions in Appendix Table 11, starting with the universe of children ever applying to a slot (column 1). In column (2) we exclude children whose parents give top ranking to private centers, and we note that imposing this restriction produces a sample where average income of the father and parental education is lower and the share of immigrant families higher. In column 3 we proceed to exclude children with priority, and note that this restriction lead to a sample where income and education is somewhat higher, in line with what we would expect given that certain priority criteria apply to disadvantaged families. Turning to column 4, we have excluded children with offers of centers that we fail to match to the registers as well as children of higher parity. The latter restriction will mechanically produce a sample with lower father's income as parents are younger (because we only include the first born child). Note that this is our lottery sample. When we compare summary statistics of column 1 and 4, we see that the lottery sample consists of families with somewhat lower average father income and parental education, and a somewhat higher share of young mothers. The share of immigrants is similar. All in all, average background characteristics are not very different across samples.

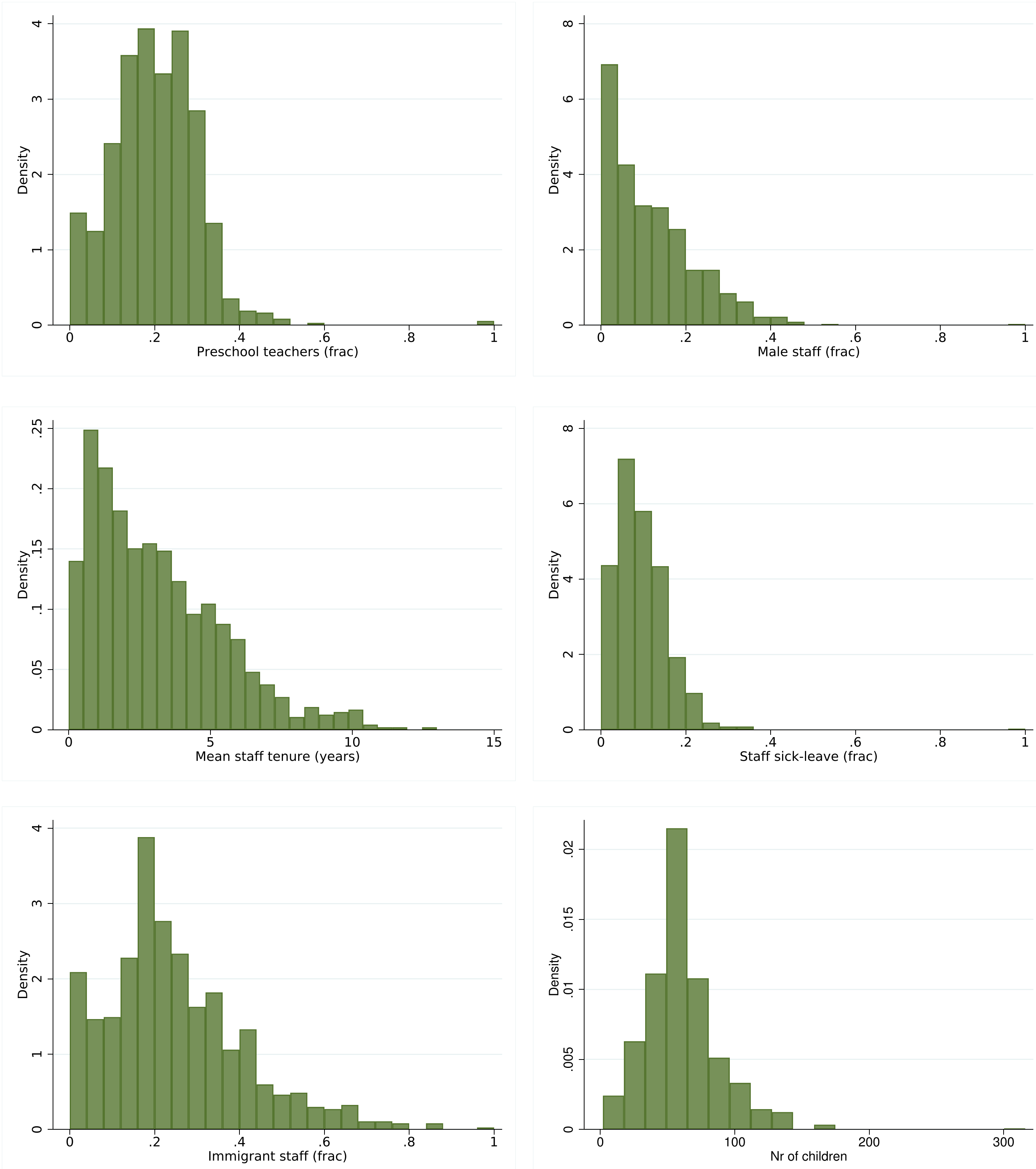

Fig. 1. Distribution of child care center characteristics. Note: N = 921 observations at the center-year level. Since observable characteristics of any center may vary in the course of a year, the unit of observation is center-year.

home to center and average share of children with immigrant background in the center. In the last row of the table, we report the mean characteristics of the first center offered for purposes of comparison. We keep in mind that the ranking in itself should not be important for the allocation process, except for the first choice center, for which families get two draws (as described in Section 2.3). However, while we have obtained information about how the allocation mechanism worked from the municipality and relevant official documents, it is not clear whether parents knew the details. Most parents were probably aware of the particular status of the first choice center, and of the priorities in their city district. Importantly, during the years we consider, information about child care center characteristics was not easily

Table 2
Characteristics of ranked centers.

Rank	Share of staff						Number of		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Male	Child care teachers	Immigrant	Sick-days	Mean staff tenure	Immigrant peers	Children/employee	Children	Distance from home (meters)
1st	0.110	0.212	0.243	0.096	3.46	0.292	3.44	68.4	901
	[0.044]	[0.000]	[0.007]	[0.567]	[0.000]	[0.000]	[0.000]	[0.274]	[0.000]
2nd	0.109	0.213	0.238	0.093	3.52	0.262	3.46	62.2	1079
	[0.094]	[0.000]	[0.001]	[0.896]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
3rd	0.118	0.209	0.243	0.096	3.56	0.258	3.61	60.7	1155
	[0.520]	[0.001]	[0.026]	[0.317]	[0.001]	[0.000]	[0.000]	[0.000]	[0.011]
4th	0.118	0.207	0.245	0.097	3.45	0.237	3.62	58.6	1176
	[0.883]	[0.210]	[0.779]	[0.044]	[0.000]	[0.000]	[0.000]	[0.000]	[0.005]
5th	0.124	0.211	0.239	0.095	3.43	0.220	3.58	57.6	1276
	[0.611]	[0.194]	[0.481]	[0.707]	[0.000]	[0.000]	[0.000]	[0.000]	[0.068]
6th	0.121	0.212	0.242	0.095	3.45	0.220	3.49	56.9	1370
	[0.3869]	[0.082]	[0.866]	[0.747]	[0.000]	[0.000]	[0.000]	[0.000]	[0.3402]
7th	0.123	0.216	0.236	0.092	3.34	0.201	3.51	57.5	1366
	[0.273]	[0.056]	[0.421]	[0.647]	[0.000]	[0.000]	[0.000]	[0.000]	[0.125]
1st offered	0.114	0.199	0.252	0.095	2.98	0.314	3.27	69.2	1325

Note: Tenure is reported in years. The unit of observation is the children. Note that a minority of the parents have ranked as many as 7 centers, and hence the number of observations differ across the rows in the table, ranging from 786 observations in the row which document characteristics of the 7th ranked center to 2175 in the last row which document characteristics of the 1st offered center.
In [ ] we report p-values from a t-test on whether the mean characteristics of the first offered center differ from the first, second etc. ranked centers.

available without getting in contact with the individual center. The cost associated with time spent on obtaining information about the characteristics listed below, was thus potentially quite high.

Even if families obtained information about structural inputs in the child care center, for example by visiting particular centers or talking to neighbors and friends, it is not clear how we should expect them to value these inputs. One point of departure may be to consider the information presented as key statistics about child care centers available today on the website of the Oslo Municipality.[17] The website contains information on size, indoor play space, number of children per employee and the share of staff with a child care teacher degree. These features may have been considered informative about the quality of a center 10 years ago as well, even though they were not available online at the time when the parents in the sample we rely on submitted applications. We note that figures for share of male staff, share of immigrant staff, sick leave and tenure are not included in today's public information about the child care centers.

From Table 2 we see that, on average, observed characteristics do not differ much across the ranked centers.[18] This is in line with what we would expect, given that ranking in itself should not matter. However, parents seem to rank centers with a higher share of immigrant peers on first rank (column 4). Also, higher ranked centers are somewhat bigger, as seen in the second to last column. While the average size of a first choice center is about 68 children, the average size of the seventh choice center is about 57. Parents also seem to care about travel time when ranking centers. In the last column, we see that travel time (in meters) increases down the ranking list. First choice centers are located on average about 900 m from the home, whereas seventh choice centers are located about 1360 m away.

In the last row, the average staff characteristics of the offered center are presented. The coefficients in brackets in the table are p-values of a test of whether the characteristics of the first ranked and the first offered center differ (row 2), whether the second ranked and first offered center differ etc. In magnitude, characteristics of the first offered center resemble those of the ranked centers to a large extent, albeit quite a few of the differences are statistically significant. For instance mean staff tenure and the fraction of staff with a child care teacher degree differ significantly across the first offered center and the first ranked. Looking at lower ranked centers, we see that the share of immigrant peers are higher in the first offered center compared to any ranked centers, and that first offered centers are bigger.[19]

We proceed to explore whether family background influences the child care center characteristics that families value. This is reported in Table 3, where we present coefficients from models in which we regress center characteristics (average tenure, share of staff with child care teacher education, share of male staff, share of staff with immigrant background, share of staff on sick leave, distance to child care center, size and child/employee ratio) of the first ranked center on the background characteristics we include in our analysis (child's gender, parental education, immigrant background and log of father's income). Families in which parental education and father's income is higher seem to appreciate a higher density of child care teachers (upper right panel), indicating a preference for more highly qualified staff. When the father has a higher income, the families seem to prefer centers with experienced employees (middle left panel), whereas they are less likely to list a center with a high share of immigrants among the staff. However, they also seem to be less concerned about bigger group size, as income is positively correlated with a 1st ranked center with a higher child/employee ratio. The share of male staff is not systematically correlated with our background variables, whereas the share of sick leave among staff is negatively correlated (at the 10% level) with both immigrant background and having parents with higher education. In the middle left panel, we see the rather puzzling pattern of families with boys being more likely to apply for child care centers with staff with long tenure compared to families with girls.

Evidence from the US suggests that the racial and ethnic composition of children attending a center is strongly correlated with the characteristics of job seekers that are invited to an interview, and indicates that this may be due to teacher hiring being influenced by customer discrimination (Boyd-Swan and Herbst, 2017). We see no indication in our data that families with immigrant background prefer child care centers with a higher share of immigrant staff. Turning to the bottom right panel, we see that higher education and income is negatively related to the share of immigrant peers in the center on first rank, whereas immigrant families prefer centers with a high share of immigrant peers.

[17] See https://www.oslo.kommune.no/barnehage/finn-barnehage-i-oslo/.
[18] We have tested whether characteristics of the first- and second-ranked center differ significantly, and find that this is the case for the share of immigrant peers, center size and distance.
[19] Note that the number of ranked centers vary across families. For instance, only 786 families have ranked 7 centers. This implies that the number of observations differ across the rows in Table 2.

Table 3
Relationship between characteristics of 1st ranked center and child background.

	Characteristics of 1st ranked center					
Child background	Share of male staff		Share of pre-school teachers		Staff sick days (share)	
Parents' educ high	0.004	(0.005)	0.009	(0.005)*	−0.005	(0.003)*
Ln(father's income)	0.001	(0.003)	0.007	(0.003)**	0.001	(0.002)
Imm. background	−0.006	(0.007)	−0.003	(0.007)	−0.010	(0.006)*
Boy	0.002	(0.004)	0.004	(0.004)	−0.001	(0.003)
	Mean staff tenure (years)		Nr of enrolled children		Distance from home (meters)	
Parents' educ high	0.027	(0.119)	−1.646	(1.360)	−40.022	(94.825)
Ln(father's income)	0.133	(0.073)*	0.286	(0.938)	−47.571	(59.872)
Imm. background	0.140	(0.183)	−1.068	(2.823)	34.893	(165.058)
Boy	0.247	(0.097)**	0.825	(1.191)	57.008	(82.984)
	Immigrant staff (share)		Nr of children/nr of employees		Immigrant peers (share)	
Parents' educ high	−0.005	(0.007)	0.016	(0.047)	−0.027	(0.008)***
Ln(father's income)	−0.018	(0.005)***	0.058	(0.033)*	−0.021	(0.006)***
Imm. background	0.003	(0.012)	−0.006	(0.082)	0.078	(0.017)***
Boy	−0.003	(0.006)	−0.041	(0.041)	0.012	(0.008)

Note: N = 2175 children. The models are estimated by OLS. Each square represents a separate regression, in total nine different specifications are reported. All dependent variables are measured at the child care center level. Included in all specifications are a constant term and dummy variables for the child's birth year and month, application year, number of ranked centers, city districts, mother's continent of origin and whether the mother and father were young (22 years or younger) when they became parents. 'Parents high educ' = a dummy variable for whether average years of schooling of mother and father is 17 or larger. 'Imm. background' is a dummy variable taking the value 1 if the child has two foreign-born parents. */**/*** denote statistical significance at the 10/5/1% level.

The latter can be explained by the fact that immigrants typically cluster in certain areas in the city (Drange and Telle, 2018), and hence apply for the same child care centers as they prefer child care close to their home.[20]

4. Empirical strategy

If parents who provide a home environment promoting positive child development are more likely to recognize child care centers that do the same, estimates of quality inputs will be biased upwards. We therefore need to take into account parental preferences for child care centers in order to get closer to the causal interpretation of center quality. As previously described, child care centers in Oslo were oversubscribed throughout the period our data covers, and slots were allocated in a lottery. As seen in Table 1, the majority of children were not admitted to their most preferred center. The random nature of slot allocation allows us to compare the outcomes of children whose parents had similar preferences, but who received offers from different child care centers due to the outcome of the lottery. In order to identify whether test scores in the first, second and third grade differ among children who, as a result of the lottery, received offers of enrollment from centers with varying staff composition, we estimate the following equation:

$$Y_{ijt} = \alpha + \beta quality_{jt} + \delta preferences_{it} + X_i + \varepsilon_i \qquad (1)$$

Y_{ijt} is the average test score for child i in the first, second and third grade whose first offered child care center was child care center j in year t. $quality_{jt}$ is a vector of quality aspects of the first child care center the child received an offer from in year t. As a measure of quality we will include the share of educated child care teachers, mean staff tenure, share of male employees, share of staff with immigrant background and average share of sickness absence (certified by a GP) among staff.

$preferences_i$ is a lottery-specific choice set included to account for the fact that parents have different preferences regarding child care centers and apply to different institutions in year t with varying characteristics and oversubscription rates. The controls for parental preferences are collected from the first application ever submitted by the parents, and we construct a choice set where all child care centers are included as separate dummies that take the value 1 if that particular center was one of the ranked centers in the application form, and 0 if it was not. Hence for children whose parents ranked seven centers (48%) there will be seven child care center dummies with the value 1. If parents ranked three centers, only three of the child care dummies take the value one. We also include dummies for how many child care centers parents have listed to account for the possible strategic behavior among families that listed less than seven centers. Moreover, as we recall from Section 2.3, families may get an extra draw in their top-ranked child care center after first receiving a lower-ranked offer. Thus, listing institutions that are expected to have low oversubscription as top-ranked will increase the likelihood of receiving an offer from this particular institution. To account for such possible strategic behavior, we control separately for the first choice center.[21]

Finally, X_i is a vector of covariates measured the year before the child is born, as well as year and cohort fixed effects described in detail in Section 3.1. As the date of birth became predictive of a child's lottery number at the end of the period, we include month of birth fixed effects to take into account possible timing of births.[22] ε_i is a random error term. Standard errors are clustered at the level of first offered child care center.

According to the municipal administration, the lottery was randomized by means of a computer algorithm. However, there is always the possibility that the randomization failed, or that manipulation occurred between the randomization and the distribution of offer letters. To investigate whether we can trust the randomization of child care offers, we turn to Table 4 where we regress center characteristics of the first offered center on child background and characteristics. All models include the same controls as in the main specification. Overall, there is little to

[20] This paper focuses on structural quality at the child care center level, and neither distance nor the share of immigrant peers in the center will be included in the main regressions. However, we do report results from specifications where distance is included (see Appendix Tables 15 and 16, Model 5) and 11), and we include peers with immigrant background in Appendix Table 13.

[21] We conducted a number of robustness tests to ensure that our specifications are robust to the choice set. While precision levels do vary, depending on how much we demand from the choice set, results are in general very similar across specifications. The results of estimations with other definitions of the choice set are available from the authors upon request.

[22] In our sample this is only relevant for a small share of children applying in 2009 and 2010. Our results are robust to exclusion of these observations. These results are available from the authors upon request.

Table 4
Relationship between center characteristics of the first center offered and child background.

	Share of staff							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Male	Childcare teachers	Immigrants	Sick-days	Staff tenure (mean)	Nr children/nr employee	Ln (nr children)	Distance from home (meters)
Ln(father's income)	−0.000	0.003	−0.009	−0.002	0.001	−0.000	−0.014	124.3
	(0.005)	(0.005)	(0.008)	(0.003)	(0.092)	(0.045)	(0.022)	(86.0)
Parents educ high	0.005	0.009	−0.018	−0.004	0.107	−0.061	0.046	32.1
	(0.007)	(0.007)	(0.011)	(0.004)	(0.144)	(0.071)	(0.034)	(148.1)
Imm background	−0.011	−0.000	0.002	0.005	0.033	-0.212^{**}	-0.082^{*}	89.7
	(0.010)	(0.010)	(0.016)	(0.006)	(0.223)	(0.102)	(0.048)	(202.3)
Boy	0.002	−0.003	0.010	−0.001	0.034	−0.084	-0.046^{*}	-372.0^{***}
	(0.005)	(0.005)	(0.008)	(0.003)	(0.121)	(0.053)	(0.027)	(109.8)

Note: N = 2175 children. Tenure is measured in years. The models are estimated by OLS. Each column represents a separate regression, in total eight different specifications are reported. The dependent variables are measured at the child care center level. Included in all specifications are dummy variables for parent's preferences where we control separately for the first ranked center, a constant term and dummy variables for the child's birth year and month, mother's continent of origin and whether the mother and father were young (22 years or younger) when they became parents, number of ranked centers, city districts, application year and interaction terms between city districts and application year. ‘Parents high educ’ = a dummy variable for whether average years of schooling of mother and father is 17 or larger. ‘Imm. background’ is a dummy variable taking the value 1 if the child has two foreign-born parents. $^{*}/^{**}/^{***}$ denotes statistical significance at the 10/5/1% level.

suggest that resourceful parents are more likely to receive an offer from a child care center of higher quality. However, and somewhat surprisingly, we see from column (7) that families with boys are somewhat more likely to get an offer from a child care center closer to their house. There is also a negative relationship between the child/employee ratio and being a boy, and a similar relationship between the ratio and having immigrant background. Both these relationships are, however, imprecise. Given the number of coefficients we test in the table, a couple of significant estimates are likely to appear due to chance. All in all we find little reason to worry that the randomization is compromised.

We also estimate whether the probability of receiving an offer from a higher ranked center depends on family background. If the allocation of children to child care centers (conditional on their preferred choices) is random, there should not be any systematic relationship between the child's background characteristics and the rank of the center from which the child received an offer. We generate an ordinal variable taking the highest value 6 if the child receives an offer from the first ranked center, and the values 5–2 if the child receives an offer from the second, third, fourth or fifth-seventh ranked center, respectively. If the child receives an offer from a center that was not in their choice set, the variable takes the value 1. The estimated relationships are presented in Table 5. When not taking account of parents' preferences for child care (column 1) we see that there is a negative and significant relationship between parents' education and receiving an offer from a higher ranked center. This may be due to parents with higher education assigning a high ranking to popular child care centers. Reassuringly, when controlling for preferences (column 2), we find no support for there being a significant relationship between observable family characteristics and receiving an offer from a higher ranked center. This underlines the importance of taking parental application behavior into account.

Table 5
The relationship between family background and the probability of being assigned to a higher ranked center, by OLS.

	(1)	(2)
Boy	0.096	0.126
	(0.083)	(0.113)
Parents educ high	-0.341^{***}	−0.096
	(0.115)	(0.148)
Imm background	0.155	−0.207
	(0.177)	(0.224)
Ln(father's income)	0.013	−0.055
	(0.069)	(0.097)
Choice set, 1, 2–7	No	Yes

Note: N = 2175 children. The models are estimated by OLS. Each column represent a separate regression. The dependent variable is an ordinal variable taking the highest value 6 if the child gets an offer from the first ranked center, and the values 5–2 if the child gets an offer from the second, third, fourth or 5th–7th ranked center. Included in all specifications are a constant term and dummy variables for the child's birth year and month, mother's continent of origin and whether the mother and father were young (22 years or younger) when they became parents, number of ranked centers, city districts, application year and interaction terms between city districts and application year. In column (2) we include dummy variables for parents' preferences and control separately for the first-ranked center. ‘Parents high educ’ = a dummy variable for whether average years of schooling of mother and father is 17 or larger. ‘Imm. background’ is a dummy variable taking the value 1 if the child has two foreign-born parents. Standard errors are clustered at the (first offered) center level and are robust to heteroskedasticity $^{*}/^{**}/^{***}$ denotes statistical significance at the 10/5/1% level.

Parents may reject their first offer for several reasons. Perhaps they have changed their minds about child care start for the child, or they have decided to move. If parents are sufficiently unhappy with the center their child was admitted to, they might decline the slot and apply again. Since unobserved parental characteristics are likely to determine who rejects their first offer and continues to search for a better option, we focus on characteristics of the child care center in which the child was initially admitted (and not the first he/she started in). This means that some of the children in our sample never start in the center we record them in. While this is not a threat to the validity of our empirical strategy, it may have implications for how we interpret the estimates. Reassuringly, a closer look at our sample reveals that as many as 87% of the parents of children in our sample are compliers, and accept the slot in the first offered center. 59% stay on in their offered center throughout the time we observe them (in most cases until school start). If deniers are families that receive an offer in a low quality institution, and these families go on to obtain a slot in a more preferred institution, it will presumably create more noise in our estimations, since some of the children we record as being in a low quality institution, instead enroll in an institution of higher quality. Ultimately, our estimates should be interpreted as the (reduced form) effect of being offered a child care slot in a center with a certain quality. Reassuringly, the majority of children accept that offer, and stay on in the center.[23]

Finally, measuring the quality of an institution is not straightforward. We cannot, for instance, rule out that some of the staff composition measures such as the share of child care teachers or the share of male staff correlate with other unobserved features of a particular center. We go on to estimate models where we include center fixed-effects to account for unobservable center characteristics such as management, premises or a specific pedagogical approach. We also perform robustness checks where we control for other observed features of a particular center. One such feature is characteristics of the peers such as the share

[23] While it would be somewhat easier to interpret our results in an IV framework instrumenting the center characteristics of the attended center with the characteristics of the first offered center, we refrain from this as it can be argued that the exclusion restriction may not hold. Some families may decline the first offer if they are unhappy with the lottery assignment, and hence postpone child care start (and starting age). If starting child care later affects child outcomes (as is found in our setting by Drange and Havnes (2019)), an IV analysis will yield biased estimates.

of children with immigrant background and the share of children with parents with higher education.

5. Results

5.1. Main results

Table 6 reports our main results. We show estimates for the relationship between child care center input and child care characteristics for language (first four columns) and mathematics (last four columns). All models are based on Eq. (1), and include controls for parental preferences for child care centers as well as for background characteristics. The first column under each subject reports the result as a percentage of the standard deviation. From the first row (starting with language results in the first column) we see that estimates for tenure are close to zero.[24] In line with the study by Goertz et al. (2018) showing positive fixed effect estimates of male staff in the child care center on child test scores, we also find that the share of males on the staff predicts positive child cognitive development. In our case, point estimates are positive for both language and mathematics, although the mathematics results are smaller and not significant. We find little evidence that test scores are higher if a child is offered a slot in a center with a higher share of staff with child care teacher degrees. This is in line with evidence from the education literature, where formal qualifications explain little of the variation in teacher quality (Krueger, 1999; Rivkin et al., 2005). The share of staff with immigrant backgrounds does not seem to cause variation in child outcomes. Interestingly, we find that children who are offered a slot in a center where employees have a higher level of sick leave, perform worse on tests in both language and mathematics during their early school years. This is similar to what has been found for school children (Clotfelter et al., 2009), for whom teacher absences have a negative effect on learning. Lastly, estimates for center size and the share of children per adults are close to zero for all outcomes.

Mean estimates may mask important heterogeneity in cases where the composition of the child care center staff is more or less important for children depending on their ability and/or family background. While keeping in mind that the distribution of the test scores is skewed, and quite different from the often bell-shaped test scores considered in the literature, we want to consider how estimates vary across the test score distribution. The remaining columns in Table 6 report results from Eq. (1) with the outcome being a dummy variable equal to one if pupil i scores above the 25th, 50th and 75th percentiles respectively.[25] Estimates reveal a pattern similar to that for the main results. In general, neither tenure, share with child care teacher education or share of staff with immigrant background turn out to be important for child cognitive development in any part of the test score distribution. However, the relationship between the share of male staff and language development is substantial and driven by children scoring above the 25th and the 50th percentiles. Sick leave among the child care staff appears to be detrimental to child achievement in the higher part of the test score distribution, for both language and mathematics. As already mentioned, the tests are intended to detect whether children are underperforming. Thus scoring above the 75th percentile may not be very informative. In some of the remaining tables we therefore only report results for whether or not the child scores above the 25th and the 50th percentiles, as we consider these two outcomes to be most relevant, given the nature of the test and consequently the distribution of the test scores.

Taking a closer look at the magnitude of increasing the share of male staff, the size of the point estimate suggests that if a child is offered a slot in a center where the share of male staff is 1 rather than 0, child language test scores increase by 52% of a standard deviation and the likelihood of scoring above the 25th and 50th percentiles increases by about 37 and 53 percentage points, respectively. While these estimates seem large, being offered a slot in a center where the share of male staff is 1 rather than 0 is a rather hypothetical exercise. When we consider Appendix Table 10 and Fig. 1, we see that virtually none of the centers have a share of male staff higher than 0.5. However, the standard deviation of the share of male staff is 0.10. Thus, receiving an offer of enrollment in a center with a one standard deviation higher share of male staff implies that a child's language test score increases by 5.2% of a standard deviation, and the likelihood of the child scoring above the 50 percentile increases by about 5.3 percentage points. The former is slightly higher, but comparable to findings in the study by Goertz et al. (2018), which shows that increasing the male share of staff by 10% translates into an increase of 3.5 and 2.1% of a standard deviation for boys and girls respectively. Similarly, being offered a slot in a center where the share of sick leave days is one standard deviation higher ($SD = 0.06$), translates into a reduction in language test scores of 17.9% of the standard deviation and a reduction in the likelihood of scoring above the 75th percentile by about 8 percentage points for language and 9.9 percentage points for mathematics. As discussed in Section 4, the results must be interpreted as the (reduced form) effect of receiving an offer from a center with a certain quality.[26]

The distribution of the center quality characteristics is somewhat skewed, and we take this into account by also estimating models where the center characteristics are defined as dummy variables equal to one if the particular characteristic at the center level is equal to or above the median or the 75th percentile in the sample of child care centers in our data. Results are reported as figures for the sake of simplicity. Results for language are displayed in Appendix Fig. 3, and results for mathematics in Appendix Fig. 4. Each sub-figure presents the results of two separate regressions. The squared (green) line represents point estimates from specifications where the center characteristics are dummy variables for being equal to or above the 50th percentile, whereas the circle (yellow) line represents point estimates from specifications where the center characteristics are dummy variables for being equal to or above the 75th percentile.

Results for sick leave and male employees are similar to what we saw in Table 6, but only significant at (or close to) the 5% level for the sub-sample of centers where the share of male staff or sick leave is equal to or above the 75th percentile. This is particularly pronounced for sick leave, as we find no negative relationship between receiving an offer from a center with sickness absence above the median and test scores in language and mathematics. When sickness absence is above the 75th percentile, however, children score substantially lower on test scores in both subjects. These findings indicate that male staff need to be above a certain threshold to be related to child language development. Moderate sick leave among staff in a center is not something to worry about, but is detrimental to child outcomes when the institution is among the 25% centers with highest sick leave.

5.2. Heterogeneity

Evidence from the classroom suggests that male teachers are particularly important for boys (Dee, 2005, 2007), but may also have a positive impact on girls' mathematics results (Dee, 2007; Antecol et al., 2015). In the child care literature, Goertz et al. (2018) find that an increase in the share of male staff does indeed predict higher child test

[24] For robustness we also run our main regression with a variable capturing the average age of the staff in the center instead of average tenure. This does not affect the results. Results are available from the authors upon request.

[25] We also constructed alternative measures for poor performance, and the results are stable.

[26] We have also produced results from a specification where we instrument for characteristics of the first child care center attended with characteristics of the first offered center. The findings are very similar, and are available from the authors upon request. 13% of the children in our sample did not accept the first offer. Even though it is endogenous whether one accepts the first offer or not, we run a regression where we exclude these children. As we can see in row (6) and (12) in Appendix Tables 15 and 16, this leaves the results unaltered.

Table 6
Main results: child care center characteristics and performance in language and mathematics.

	Language				Mathematics			
	% of SD	≥25th	≥50th	≥75th	% of SD	≥25th	≥50th	≥75th
Tenure	0.002	−0.009	0.011	0.035**	0.001	0.006	0.024	0.010
	(0.027)	(0.015)	(0.020)	(0.018)	(0.033)	(0.020)	(0.022)	(0.017)
Male	0.522**	0.367***	0.533***	0.169	0.220	0.096	0.065	−0.024
	(0.242)	(0.131)	(0.170)	(0.165)	(0.267)	(0.150)	(0.184)	(0.158)
Child care teachers	−0.228	−0.113	0.021	−0.140	0.097	0.188	0.024	0.023
	(0.274)	(0.131)	(0.180)	(0.186)	(0.257)	(0.144)	(0.182)	(0.172)
Imm background	0.008	0.050	−0.037	−0.023	0.145	0.042	0.106	0.034
	(0.211)	(0.107)	(0.139)	(0.116)	(0.209)	(0.113)	(0.133)	(0.107)
Sick-days	1.076**	−0.297	−0.544*	−0.480*	−0.778	−0.439	−0.394	−0.595**
	(0.442)	(0.233)	(0.280)	(0.280)	(0.508)	(0.279)	(0.309)	(0.266)
Ln(nr. children)	−0.016	0.025	0.012	0.005	−0.078	−0.006	−0.005	−0.015
	(0.064)	(0.034)	(0.037)	(0.036)	(0.059)	(0.031)	(0.036)	(0.031)
Nr children/nr employee	−0.014	−0.001	−0.019	−0.026	−0.016	−0.019	−0.012	−0.017
	(0.032)	(0.017)	(0.022)	(0.021)	(0.031)	(0.018)	(0.019)	(0.018)

Note: N = 2175 children. Estimates are from estimating Eq. (1) with OLS regression. All child care characteristics are measured at the child care center level. As a measure of tenure we use average years the current employees have been working in a given center. Other staff characteristics are measured as shares. Included in all specifications are dummy variables for parents' preferences where we control separately for the first-ranked center, a constant term and the family background variables listed in Section 3.1 and Appendix Table 10, (staff's) tenure squared and dummy variables for the child's birth year and month, mother's continent of origin, number of ranked centers, application year, city districts and interaction terms between application year and city districts. Standard errors are clustered at the (first offered) center level and are robust to heteroskedasticity. */**/*** denotes statistical significance at the 10/5/1% level.

scores in 2nd grade, but the results are similar for boys and girls. In Panel A of Appendix Table 12 we report results from a regression where we interact the child's gender (boy) with the fraction of male employees. The sign of the interaction coefficient varies across specifications but is not statistically significant.[27]

It has been hypothesized that teachers could be more responsive to the needs of children who share their racial or ethnic background (Dee, 2004). Dee (2004) finds for the US that both black and white children benefit from a same-race teacher, while results for Denmark in Bauchmuller et al. (2014) suggest that ethnic minority children gain significantly less from a higher share of ethnic minority staff than children without such background. In panel B of Appendix Table 12 we investigate this for Norway and display the results of interacting the child's immigrant background with the fraction of the staff with immigrant background. The sign of the interaction term is not statistically significant. In brief, overall we see little evidence of a pattern of difference across groups.

5.3. Robustness

In our main specification we present estimates from multivariate regressions that include all the center characteristics. In Table 7 we report results from separate regressions for the center characteristics. The idea is to check whether some of these characteristics are correlated. The results of these bivariate regressions are very similar to the multivariate regression in Table 6 with an exception for tenure, for which estimates are negative for language development.

In Appendix Tables 15 and 16 we present results from several specifications to further explore the robustness of our results. The lower panel of the table indicates which controls are included in the respective models. We report results for whether child test scores are above the 25th or 50th percentile, but for simplicity we will focus on the latter when discussing results. We start out with a naive specification where we regress child test scores on child care characteristics (Model 7), and go on to add background characteristics in Model 8. In Model 9 we add preferences, giving us the same specification as our main results. Comparing these estimates may give us an idea of how families select their children into centers. Point estimates for male employees are similar across specifications in model 7–9, particularly for language results. Findings for the share of staff on sick-leave are stronger when we add preferences in 9 (compared to Model 7 and 8).

In the Norwegian school admission system, children are admitted to schools depending on catchment area. While parents may apply to another school than their catchment area school, this is not common and such admission will only be granted if there are available slots after catchment area children have been admitted. Nevertheless, we cannot rule out that some families choose to apply for another school based on which child care center they were admitted to. Compared with the main results in Model 9, including school fixed effects in Model 10 leave the results unaltered.

A closer distance to the center may lead to more quality time with the parents or to more interaction with peers on weekends. We also remember from Table 2 that parents seem to prefer a center closer to the home. We go on to explore this by including distance from home as a separate variable. In row 11 in Appendix Tables 15 and 16 we see that the point estimate for distance is very small, and the other point estimates are similar to the main specification. This suggests that distance in our setting do not explain variations in child performance.

While our specification controls for selection of children into child care centers, we do not take into account the selection of staff to different child care centers. This implies that the positive point estimate of male staff and the negative point estimate of sickness absence could capture that male employees and sickness absence are correlated with observed and/or unobserved features of the center that are not picked up in our model and which also affects children's cognitive development. We go on to look at whether centers that differ in terms on male staff and staff on sick leave also differ along other observable dimensions. In Table 8 we divide child care centers into two groups depending on whether the share of male staff (panel A) and the share of staff sick-days (panel B) is below or above the median. In [] we report p-values from a t-test on whether the characteristics are significantly different across the two groups of centers. While we see that tenure is somewhat lower in centers with a high share of male employees, and that these centers also are on average bigger, other differences are small in magnitude. Turning to Panel B, we note that centers with high sickness absence have a somewhat higher share of staff with immigrant background and also here tend to be bigger. Otherwise differences here, too, are small.

Another way of exploring whether male staff and staff with better unobserved health select into "better" centers, could be to account for

[27] It has been hypothesized that male staff may be important role models for children with few male role models at home (such as children of single mothers). In order to investigate this, we estimate a model in which we interact a dummy variable that takes the value one if the mother was registered as having no partner the year the child was born with the fraction of male staff. The results give no support to this hypothesis.

Table 7
Child care center characteristics and child performance in language and mathematics, separate regressions for each outcome.

	Language				Mathematics			
	% of SD	≥25th	≥50th	≥75th	% of SD	≥25th	≥50th	≥75th
Tenure	0.031**	−0.011*	−0.016*	0.000	−0.018	−0.004	−0.002	−0.005
	(0.015)	(0.006)	(0.008)	(0.009)	(0.015)	(0.008)	(0.010)	(0.008)
Male	0.576**	0.390***	0.585***	0.179	0.219	0.121	0.055	−0.008
	(0.245)	(0.129)	(0.170)	(0.164)	(0.267)	(0.150)	(0.183)	(0.152)
Child care teachers	−0.209	−0.114	0.087	−0.071	−0.001	0.182	0.009	0.015
	(0.235)	(0.124)	(0.163)	(0.171)	(0.226)	(0.130)	(0.168)	(0.163)
Imm background	0.017	0.037	−0.093	−0.037	0.126	−0.024	0.084	0.012
	(0.192)	(0.099)	(0.129)	(0.104)	(0.192)	(0.102)	(0.121)	(0.102)
Sick-leave	−1.065**	−0.319	−0.528*	−0.390	−0.715	−0.389	−0.320	−0.536**
	(0.442)	(0.234)	(0.284)	(0.277)	(0.503)	(0.273)	(0.306)	(0.264)
Ln(nr children)	−0.003	0.032	0.025	−0.007	−0.072	−0.003	−0.013	−0.018
	(0.063)	(0.032)	(0.036)	(0.036)	(0.057)	(0.029)	(0.034)	(0.032)
Nr children/nr staff	−0.017	−0.000	−0.023	−0.025	−0.017	−0.017	−0.009	−0.013
	(0.031)	(0.016)	(0.021)	(0.020)	(0.031)	(0.017)	(0.019)	(0.017)

Note: N = 2175 children. The models are estimated by OLS, with a separate regression for each quality indicator. All child care characteristics are measured at the child care center level. As a measure of tenure we use average years the current employees have been working in a given center. Other staff characteristics are measured as shares. Included in all specifications are dummy variables for parents' preferences where we control separately for the first-ranked center, a constant term and the family background variables listed in Section 3.1 and Appendix Table 10 and dummy variables for the child's birth year and month, mother's continent of origin, number of ranked centers, application year, city districts and interaction terms between application year and city districts. In the model where we estimate the reduced form effect of (staff's average) tenure we also control for (staff's average) tenure squared. Standard errors are clustered at the (first offered) center level and are robust to heteroskedasticity. */**/*** denotes statistical significance at the 10/5/1% level.

Table 8
Characteristics of centers with low vs high share of male employees and sickness absence.

	Panel A			Panel B		
	Male staff (share)			Staff sick-days (share)		
	Low	High	Diff	Low	High	Diff
Mean staff tenure (years)	3.53	2.76	[0.000]	3.15	3.19	[0.812]
Child care teachers (share)	0.20	0.21	[0.200]	0.21	0.20	[0.304]
Male staff (share)	0.04	0.21	[0.000]	0.12	0.11	[0.474]
Imm background (share)	0.25	0.25	[0.840]	0.24	0.26	[0.027]
Staff Sick-days (share)	0.09	0.10	[0.450]	0.05	0.14	[0.000]
Nr children/nr employees	3.4	3.2	[0.002]	3.4	3.3	[0.0540]
Nr children	61	66	[0.005]	60	66	[0.002]
N	492	429		421	500	

Note: Remark about unit of observation: Since observable characteristics of any center may vary in the course of a year, the unit of observation is center-year. All child care characteristics are measured at the center-year level and reported as means. As a measure of tenure we use average years the current employees have been working in a given center. Other staff characteristics are measured as shares.
In [] we report p-values from a *t*-test on whether the mean characteristics in centers with a high share of male and a high share of sickness absence differ from centers with a low share of male employees and a low share of sickness absence.

peer characteristics. If such selection does occur among staff, controlling for the share of children with college-educated parents and the share of children with an immigrant background should lead to smaller effect estimates. Results when adding peer characteristics are reported in Appendix Table 13 and are very similar to those in the main specification in Table 6.[28]

Male and female staff may differ along dimensions that are not picked up by our main specification, for instance may the child care teacher degree fail to capture important cognitive skills. Or perhaps female and male staff differ when it comes to cognitive skills that are unobserved at the time of observation. In Table 9 we report averages of observed individual characteristics for females (column 1) and males (column 2) working in the child care centers in our sample. In the third column we report the difference in mean and the p-value from a *t*-test for whether the mean characteristics differ across females and males. Row 1 of Table 9 displays the mean years of schooling *ever* completed (also after we have observed them in the child care center) and reveals that this do not differ between male and female child care staff. On average male staff are less sick than female staff. They are also younger, and have shorter center tenure. The latter explains why tenure is lower in centers with a high share of male staff, like we observed in Table 8 above. Given that sick-leave and center tenure already is accounted for in our results, and that the point estimates for male employees are very similar in the multivariate (Table 6) and the bivariate (Table 7) regressions, it seems unlikely that these differences can explain that children in centers with a high share of male staff perform better on tests in language.

While we find little evidence that the shares of male employees and sick leave at the center level correlate systematically with observables that may explain the effect on child development, the question as to whether something unobservable correlates with the quality characteristics still remains. We go on to add center fixed effects to take into account time-invariant conditions in the center such as management, premises or a specific pedagogical approach that may also influence the children's performance. We report estimates from this specification in Appendix Table 14. Keeping in mind that this specification demands quite a lot from the data, we should not be surprised by the higher standard errors. From the first column (language) and fifth column

[28] We construct measures of peer composition by averaging the background characteristics of children from a certain cohort in a certain year enrolled in a certain center. For example, for children born in 2004, we will construct measures of peer quality by matching each center with its respective 2004-born children, and then averaging parental education and immigrant background on the basis of these children.

Table 9
Background statistics on child care staff, by gender.

	Female		Male		Difference		N
	Mean	St.dev	Mean	St.dev	Mean	p-Val	
Years of schooling	12.38	(4.23)	12.46	(3.74)	0.082	[0.424]	15,686
Long term sick leave (frac), certified by GP	0.10	(0.21)	0.06	(0.15)	−0.044	[0.000]	16,697
Tenure (years)	3.2	(5.1)	1.6	(2.9)	−1.60	[0.000]	16,036
Age	37.90	(12.01)	31.18	(10.78)	−6.73	[0.000]	16,697

Note: Unit of observation is the individual child care center employee.
In [] we report p-values from a t-test on whether the mean characteristics differ across gender.

(mathematics) we see that estimates for the male share of staff appear to be even stronger in the fixed effects specification, albeit less precisely estimated. For mathematics, too, point estimates are positive, but very imprecise. All in all it seems safe to conclude that the positive relationship between the share of male staff and child language development is robust to the inclusion of center fixed-effects, a finding that supports a causal interpretation of the estimates. Estimates for sick leave are very similar in magnitude with regard to language, but somewhat smaller for mathematics.

6. Conclusion

Well-identified causal studies of child care quality are scarce. To credibly estimate the causal effects of child care quality, we need to account for endogenous sorting of children into centers. We aim to expand this literature by assessing how parents value structural quality inputs in the child care center, and whether child cognitive development differs systematically depending on the characteristics of the child care center. We account for possible selection into centers with different characteristics by taking advantage of a unique data set on applications for child care, detailed records on use of child care and an allocation mechanism that randomly matched children to centers conditioned on observable characteristics. During the years covered by our data, child care centers in Oslo were oversubscribed. While most children who wanted to attend a child care center would eventually enroll, the majority ended up enrolling in a different child care center from the one they actually preferred. Our rich registry data allow us to study how test scores in primary school differ among children whose parents initially applied for the same center(s), but due to the lottery got offers of enrollment from centers with variations in staff education, tenure, sickness absence, as well as in their shares of male and immigrant staff and the child/employee ratio.

Our findings suggest that parents prefer child care centers that are situated close to the home, and that the likelihood of choosing a center with a high share of child care teachers increases with father's income. In line with Walters (2015), we find no indication that child test scores differ depending on teacher characteristics such as education. We do, however, find that children allocated to centers with a higher share of male staff perform better on tests in language in their early school years, with results similar to those found for Denmark (Goertz et al., 2018). Sub-sample analysis suggests that male staff may be important for the development of both boys and girls. Higher sickness absence in a center predicts lower child test scores in both language and mathematics, as is also seen for school children in the US (Clotfelter et al., 2009).

Looking more closely for possible explanations, we find little evidence that centers with a high share of male staff or of sickness absence differ along observable dimensions compared to centers with a low share. Male employees in the child care center are somewhat younger and have lower sickness absence compared to their female counterparts. However, given that models with these explanatory variables included yield similar estimates, there is not much to suggest that lower age and less sick leave can explain the male effect. Nor does including peer characteristics, such as the parental education and immigrant background of other children in the center change the estimates for either male staff or sickness absence. We go on to estimate center fixed-effects that take into account features of the center such as pedagogical approach, premises or management. This approach yields a similar estimate for the positive relationship between the share of male staff and child development, supporting a causal interpretation of the result. As regards the negative relationship between high sick leave and child development, we find similar estimates overall, though they are slightly less robust than the finding for male staff.

Several mechanisms may explain our findings. First, as suggested by Dee (2007), male teachers can influence children's engagement or behavior by acting as role models. Second, if same-gender teachers communicate different expectations to boys and girls, the presence of male staff in the child care center could be positive for child development. None of our sub sample analyses indicate that boys are affected differently from girls, perhaps suggesting that role model theory is less important in our setting. Drawing on theory and evidence from the development psychology literature, certain types of physical play which are typically initiated by males, have been shown to facilitate social skills practice and aggression regulation (Storli and Sandseter, 2017). Norwegian child care centers focus on learning through play, and in such a setting it may matter whether all activities are initiated by female staff or whether male staff also contributes. Ultimately, while our findings may indicate that male employees interact with the children in a different way from their female co-workers, we still cannot rule out that men who decide to have a career in the female dominated child care sector may be particularly motivated and/or suited for working with children, and that this enhances child development. When it comes to understanding the negative effect of high sick leave, we lean on findings from the psychology literature that emphasize the importance of stable child-caregiver relationships during early childhood for later development (Bowlby, 1969; Sroufe et al., 2010). We study a setting where children enter child care at an early age, and where the presence of stable caregivers in the center may be particularly important for early child development. Moreover, if there are fewer staff at work than originally planned, it seems self-evident that there is less time to provide a stimulating environment for the individual child.

There has long been a consensus among psychologists, neurobiologists and economists that investing in early childhood is imperative, as this is a particularly sensitive period when the child's brain is at its most receptive, and the foundation for cognitive and socio-emotional capacities is being developed (Knudsen et al., 2006). Our results suggest that public policy should take into account the gender balance of staff in early childhood education, and that managers should pay attention to high sick leave among employees in the sector. Our findings also provide valuable information for parents who have to decide whether to enroll their child in a center with a high vs a low share of male staff, or high vs low sickness absence among the staff.

Appendix

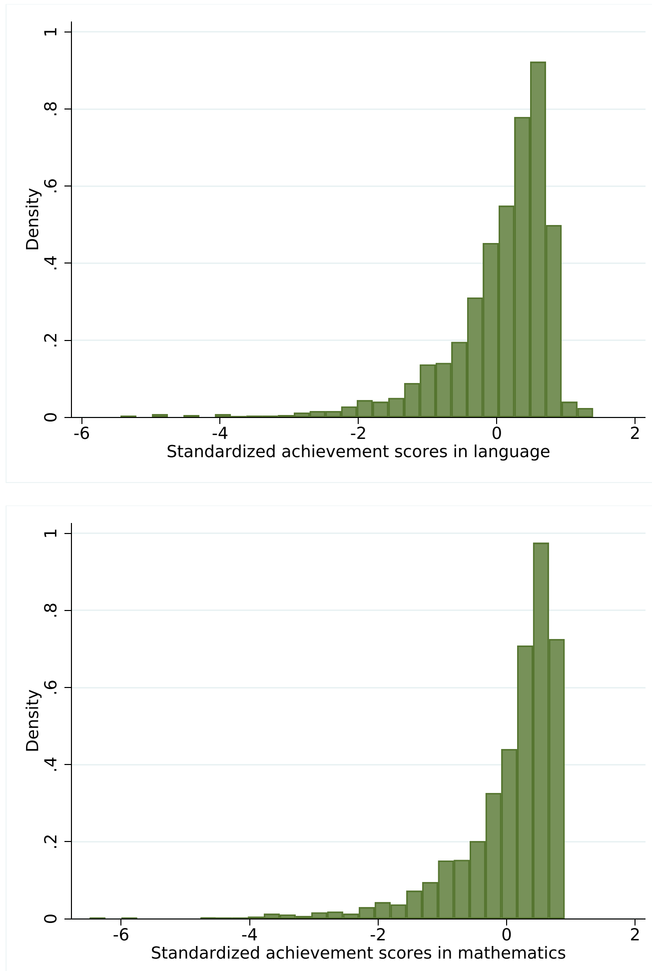

Fig. 2. Distribution of scores in language and mathematics.

Table 10
Summary statistics of all included variables.

	Mean	St.dev
Outcome variables		
Standardized achievement scores in language	0.05	0.80
Standardized achievement scores in mathematics	0.04	0.83
Explanatory variables		
Child care center level		
Tenure in years	2.98	2.36
Share of staff with child care teacher education	0.199	0.097
Share of male employees	0.115	0.102
Share of employees with an immigrant background	0.253	0.161
Share of long termsick-leave among staff	0.095	0.059
Nr. of children	67.86	33.30
Nr. of children/Nr. of employees	3.28	0.86
Individual level (reported are averages)		
Boy	0.504	0.500
Father's income (NOK)	290,324	1,374,957
Parents high education (17+)	0.38	0.49
Immigrant background	0.254	0.435
Young mother	0.077	0.266
Young father	0.025	0.157
Age when receiving the first offer	1.8	0.95

N = 2175. Long term sick-leave is all sick leave spells lasting for more than ten days. We obtain our measure of long term sick-leave at the center level by dividing the number days absent due to long-term sick-leave with number of contractual working days. Parents are defined to have higher education if average years of schooling of mother and father are 17 or larger. The child is defined to have immigrant background if it is born by two foreign-born parents. Mothers and fathers are defined to be young if they were younger than 22 years when the child was born.

Table 11
Summary statistics of background characteristics across different samples, reported are averages.

	Children ever applying	Excl private on 1 rank	Excluding Priority	Lottery sample
Boy	0.504	0.505	0.501	0.504
Father's income (NOK)	359,538	313,097	326,314	290,324
Parents high educ (17+)	0.42	0.37	0.39	0.38
Immigrant background	0.25	0.31	0.29	0.254
Young mother	0.04	0.05	0.04	0.077
Young father	0.01	0.02	0.01	0.025
Nr obs	19,281	13,724	10,611	2175

Note: The unit of observation is the child. Parents are defined to have higher education if average years of schooling of mother and father are 17 or larger. The child is defined to have immigrant background if it is born by two foreign-born parents. Mothers and fathers are defined to be young if they were younger than 22 years when the child was born.

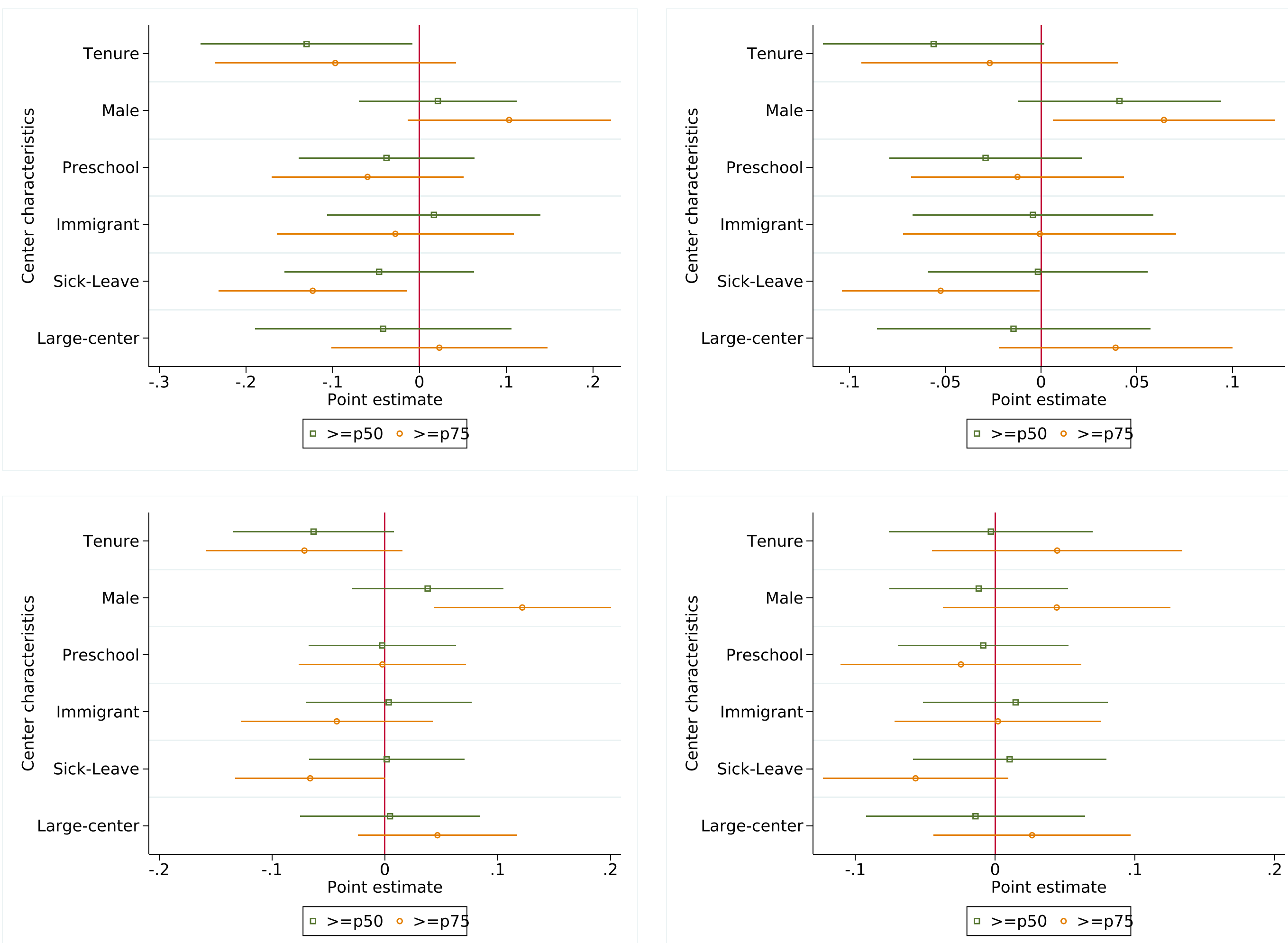

Fig. 3. Performance in reading. Note: N = 2175 children. The figure reports OLS estimates from estimating Eq. (1) when measuring center characteristics (at the center level) as dummy variables taking the value one if equal to or above the median or 75th percentile in the population of child care centers: 'Tenure' = dummy variable taking the value 1 if average tenure (in years) among staff in the particular center is equal to or above the median or 75th percentile; 'Male' = dummy variable taking the value one of the share of male staff in the particular center is equal to or above the median or the 75the percentile; 'Child care' = dummy variable taking the value one if the share of staff with a child care teacher degree in the particular center is equal to or above the median or the 75th percentile; 'Immigrant' = a dummy variable taking the value one of the share of staff in the particular center with an immigrant background is equal to or above the median or the 75th percentile; 'Sick-leave' = a dummy variable taking the value one if the share of days the staff has been absent due to long term sick-leave (spells lasting longer than 10 days) is equal to or above the median or the 75th percentile. Included in all specifications are dummy variables for parents' preferences where we control separately for the first-ranked center, a constant term and the family background variables listed in Section 3.1 and Appendix Table 10, dummy variables for the child's birth year and month, mother's continent of origin, number of ranked centers, application year, city districts and interaction terms between application year and city districts. Standard errors are clustered at the first offered institution level and are robust to heteroskedasticity.

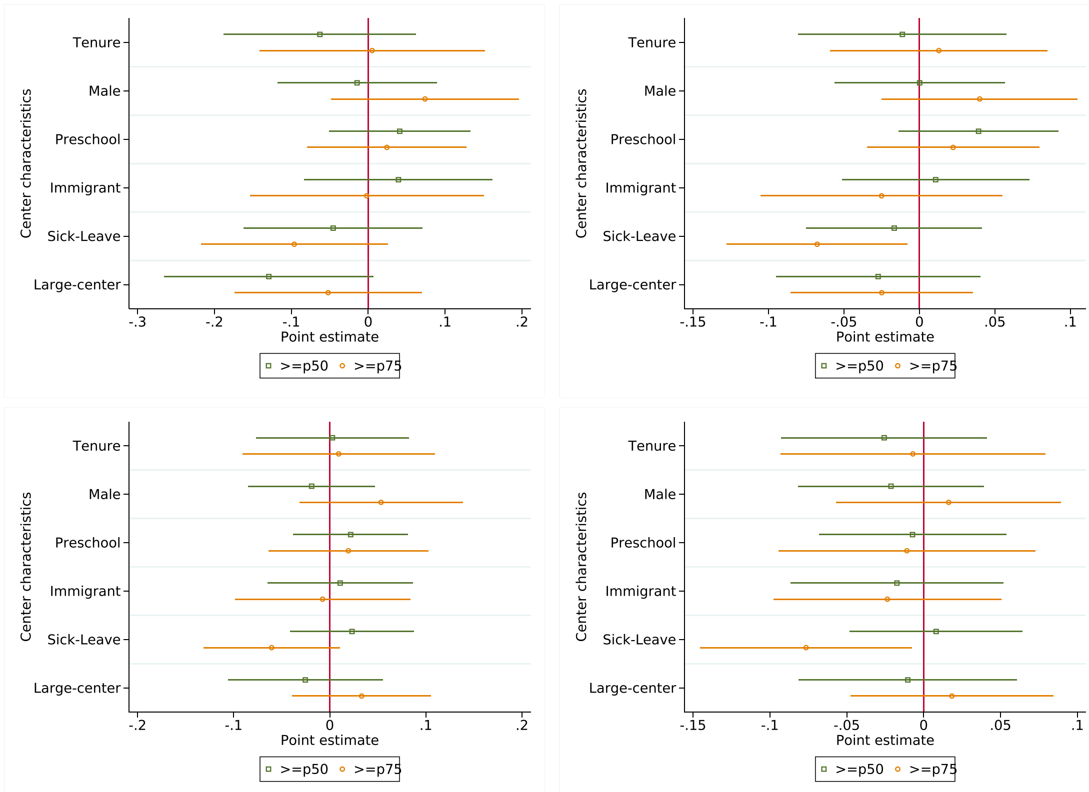

Fig. 4. Performance in mathematics. Note: N = 2175 children. The figure reports OLS estimates from estimating Eq. (1) when measuring center characteristics (at the center level) as dummy variables taking the value one if equal to or above the median or 75th percentile in the population of child care centers: 'Tenure' = dummy variable taking the value 1 if average tenure (in years) among staff in the particular center is equal to or above the median or 75th percentile; 'Male' = dummy variable taking the value one of the share of male staff in the particular center is equal to or above the median or the 75the percentile; 'Child care' = dummy variable taking the value one if the share of staff with a child care teacher degree in the particular center is equal to or above the median or the 75th percentile; 'Immigrant' = a dummy variable taking the value one of the share of staff in the particular center with an immigrant background is equal to or above the median or the 75th percentile; 'Sick-leave' = a dummy variable taking the value one if the share of days the staff has been absent due to long term sick-leave (spells lasting longer than 10 days) is equal to or above the median or the 75th percentile. Included in all specifications are dummy variables for parents' preferences where we control separately for the first-ranked center, a constant term and the family background variables listed in Section 3.1 and Appendix Table 10, dummy variables for the child's birth year and month, mother's continent of origin, number of ranked centers, application year, city districts and interaction terms between application year and city districts. Standard errors are clustered at the first offered institution level and are robust to heteroskedasticity.

Table 12
Subsample analysis: The effect of child care center characteristics on children's performance in language and mathematics, when including interaction terms.

	A				B			
	≥25th		≥50th		≥25th		≥50th	
	L	M	L	M	L	M	L	M
Tenure	−0.009	0.006	0.010	0.024	−0.010	0.006	0.012	0.025
	(0.015)	(0.020)	(0.020)	(0.023)	(0.015)	(0.020)	(0.020)	(0.022)
Male	0.300*	0.210	0.783***	0.220	0.368***	0.097	0.532***	0.065
	(0.181)	(0.200)	(0.210)	(0.229)	(0.132)	(0.150)	(0.170)	(0.184)
Child care teachers	−0.115	0.191	0.028	0.028	−0.107	0.191	0.013	0.021
	(0.132)	(0.144)	(0.181)	(0.182)	(0.133)	(0.145)	(0.181)	(0.181)
Imm background	0.049	0.043	−0.034	0.109	0.003	0.021	0.020	0.124
	(0.107)	(0.114)	(0.139)	(0.134)	(0.119)	(0.122)	(0.158)	(0.144)
Sick-days	−0.307	−0.421	−0.505*	−0.370	−0.295	−0.438	−0.547*	−0.395
	(0.234)	(0.278)	(0.275)	(0.305)	(0.233)	(0.280)	(0.281)	(0.309)
Ln(nr. children)	0.024	−0.005	0.014	−0.004	0.026	−0.006	0.011	−0.005
	(0.034)	(0.031)	(0.037)	(0.036)	(0.034)	(0.031)	(0.038)	(0.036)
Nr children/nr employee	−0.001	−0.019	−0.019	−0.012	−0.001	−0.019	−0.019	−0.012
	(0.017)	(0.018)	(0.022)	(0.019)	(0.017)	(0.018)	(0.022)	(0.019)
Boy*Male employees (frac)					Imm*Imm staff (frac)			
	0.135	−0.232	−0.506	−0.315	0.175	0.077	−0.214	−0.066
	(0.252)	(0.268)	(0.315)	(0.305)	(0.212)	(0.211)	(0.229)	(0.214)

Note: N = 2175 individuals. The table reports OLS estimates from estimating Eq. (1) with interaction terms between the child's gender and the fraction of male employees (panel A) and between the child's immigrant status and the fraction of employees with an immigrant background (panel B). L = Language, M = Mathematics. All child care characteristics are measured at the child care center level. As a measure of tenure we use average years the current employees have been working in a given center. Other staff characteristics are measured as shares. Included in all specifications are dummy variables for parents' preferences where we control separately for the first-ranked center, a constant term and the family background variables listed in Section 3.1 and Appendix Table 10, (staff's) tenure squared and dummy variables for the child's birth year and month, mother's continent of origin, number of ranked centers, application year, city districts and interaction terms between application year and city districts. Standard errors are clustered at the (first offered) center level and are robust to heteroskedasticity. */**/*** denotes statistical significance at the 10/5/1% level.

Table 13
Mechanisms: The effect of child care center characteristics on children's performance in language and mathematics, when controlling for peer characteristics.

	Language		Mathematics	
	≥25th	≥50th	≥25th	≥50th
Tenure	−0.010	0.012	0.007	0.025
	(0.015)	(0.020)	(0.019)	(0.022)
Male	0.370***	0.529***	0.104	0.042
	(0.132)	(0.172)	(0.152)	(0.186)
Child care teachers	−0.101	0.014	0.181	−0.024
	(0.134)	(0.185)	(0.147)	(0.186)
Imm. background	0.037	−0.027	0.040	0.170
	(0.109)	(0.140)	(0.114)	(0.141)
Sick-days	−0.309	−0.569**	−0.449	−0.348
	(0.233)	(0.281)	(0.277)	(0.316)
Ln(nr children)	0.024	0.012	−0.007	0.002
	(0.034)	(0.038)	(0.031)	(0.037)
Nr children/nr employees	−0.003	−0.018	−0.019	−0.004
	(0.017)	(0.022)	(0.018)	(0.020)
Peer characteristics				
College edu	0.002	−0.109	−0.149	0.063
parents(share)	(0.114)	(0.144)	(0.118)	(0.136)
Immigrant	0.039	−0.112	−0.064	−0.168
background (share)	(0.097)	(0.116)	(0.102)	(0.114)

Note: N = 2175 individuals. The table reports OLS estimates from estimating Eq. (1) when controlling for peer characteristics. All child care characteristics are measured at the child care center level. As a measure of tenure we use average years the current employees have been working in a given center. Other staff characteristics are measured as shares.
Included in all specifications are dummy variables for parents' preferences where we control separately for the first-ranked center, a constant term and the family background variables listed in Section 3.1 and Appendix Table 10, (staff's) tenure squared and dummy variables for the child's birth year and month, mother's continent of origin, number of ranked centers, application year, city districts and interaction terms between application year and city districts. Standard errors are clustered at the first offered institution level and are robust to heteroskedasticity. */**/*** denotes statistical significance at the 10/5/1% level.

Table 14
The effect of child care center characteristics on children's performance in language and mathematics, when controlling center fixed effects.

	Language				Mathematics			
	% of SD	≥25th	≥50th	≥75th	% of SD	≥25th	≥50th	≥75th
Tenure	−0.019	−0.006	0.048	0.070	−0.074	−0.076	0.001	−0.030
	(0.112)	(0.055)	(0.079)	(0.068)	(0.103)	(0.067)	(0.076)	(0.065)
Male	1.567**	1.071***	0.573	−0.142	0.859	0.290	0.671	0.249
	(0.723)	(0.354)	(0.440)	(0.436)	(0.776)	(0.412)	(0.477)	(0.401)
Child careteachers	0.543	0.255	0.060	0.108	0.932	0.404	0.255	0.595
	(0.628)	(0.330)	(0.421)	(0.422)	(0.698)	(0.365)	(0.453)	(0.465)
Imm background	0.280	0.254	0.044	0.026	0.351	−0.134	0.121	0.316
	(0.609)	(0.342)	(0.424)	(0.335)	(0.642)	(0.400)	(0.374)	(0.333)
Sick-days	−1.161	−0.245	−0.472	−0.535	−0.066	0.088	−0.323	−0.272
	(0.918)	(0.464)	(0.581)	(0.553)	(0.986)	(0.491)	(0.563)	(0.495)
Ln(nr children)	−0.322	−0.120	−0.114	0.026	−0.278	−0.117	−0.256*	0.020
	(0.218)	(0.122)	(0.142)	(0.184)	(0.233)	(0.128)	(0.139)	(0.105)
Nr children/nr employee	0.010	0.010	−0.008	−0.057	0.010	0.001	−0.001	0.004
	(0.068)	(0.034)	(0.051)	(0.044)	(0.069)	(0.042)	(0.044)	(0.053)

Note: N = 2175 children. The table reports OLS estimates from estimating Eq. (1) with (first offered) center fixed effects. All child care characteristics are measured at the child care center level. As a measure of tenure we use average years the current employees have been working in a given center. Other staff characteristics are measured as shares.
Included in all specifications are dummy variables for parents' preferences where we control separately for the first ranked center, a constant term and the family background variables listed in Section 3.1 and Appendix Table 10, (staff's) tenure squared and dummy variables for the child's birth year and month, mother's continent of origin, number of ranked centers, application year, city districts and interaction terms between application year and city districts. Standard errors are clustered at the first offered institution level and are robust to heteroskedasticity. */**/*** denotes statistical significance at the 10/5/1% level.

Table 15
Robustness checks for language.

	≥25th						≥50th					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Tenure	0.007	0.007	−0.009	−0.007	−0.011	−0.006	0.001	−0.001	0.011	0.019	0.011	0.025
	(0.010)	(0.010)	(0.015)	(0.016)	(0.016)	(0.021)	(0.011)	(0.011)	(0.020)	(0.021)	(0.021)	(0.026)
Male	0.289***	0.279***	0.367***	0.389**	0.375***	0.339**	0.385***	0.370***	0.533***	0.534***	0.478**	0.547***
	(0.080)	(0.079)	(0.131)	(0.155)	(0.138)	(0.167)	(0.109)	(0.106)	(0.170)	(0.179)	(0.188)	(0.197)
Child care teachers	−0.001	−0.028	−0.113	−0.183	−0.086	−0.110	0.037	0.004	0.021	−0.023	0.019	−0.089
	(0.090)	(0.090)	(0.131)	(0.142)	(0.139)	(0.168)	(0.125)	(0.122)	(0.180)	(0.182)	(0.188)	(0.205)
Imm background	0.006	0.022	0.050	0.043	0.045	0.047	0.003	0.024	−0.037	−0.027	−0.038	−0.100
	(0.059)	(0.058)	(0.107)	(0.110)	(0.107)	(0.141)	(0.075)	(0.074)	(0.139)	(0.133)	(0.140)	(0.172)
Sick-days	−0.166	−0.123	−0.297	−0.208	−0.288	−0.140	−0.089	−0.076	−0.544*	−0.464	−0.543*	−0.468
	(0.139)	(0.134)	(0.233)	(0.253)	(0.247)	(0.255)	(0.190)	(0.180)	(0.280)	(0.285)	(0.301)	(0.330)
Ln(nt children)	0.005	0.004	0.025	0.021	0.017	−0.010	−0.000	−0.002	0.012	0.011	0.012	−0.003

(continued on next page)

Table 15 (*continued*)

	≥25th						≥50th					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	(0.016)	(0.016)	(0.034)	(0.034)	(0.035)	(0.039)	(0.023)	(0.023)	(0.037)	(0.038)	(0.039)	(0.046)
Nr children/nr employee	0.003	0.004	−0.001	0.000	0.001	0.011	−0.021*	−0.021*	−0.019	−0.020	−0.021	−0.006
	(0.010)	(0.009)	(0.017)	(0.018)	(0.018)	(0.021)	(0.012)	(0.012)	(0.022)	(0.022)	(0.023)	(0.024)
Distance					−0.001						−0.003	
					(0.008)						(0.009)	
Controlling for:												
- Fam char	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
- Choice set	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
- School FE	No	No	No	Yes	No	No	No	No	No	Yes	No	No
Keep those who accepted 1st offer						Yes						Yes
Nr of obs	2175	2175			2121	1896	2175	2175			2121	1896

Note: The table reports OLS estimates from estimating different variants of Eq. (1) for language.. All child care characteristics are measured at the child care center level. As a measure of tenure we use average years the current employees have been working in a given center. Other staff characteristics are measured as shares.
Included in all specifications are dummy variables for parent's preferences where we control separately for the first ranked center, a constant term and the family background variables listed in Section 3.1 and Appendix Table 10, (staff's) tenure squared and dummy variables for the child's birth year and month, mother's continent of origin, number of ranked centers, application year, city districts and interaction terms between application year and city districts. Standard errors are clustered at the (first offered) center level and are robust to heteroskedasticity. */**/*** denotes statistical significance at the 10/5/1% level.

Table 16
Robustness checks for mathematics.

	≥25th						≥50th					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Tenure	0.004	0.004	0.006	0.013	0.008	0.020	0.020*	0.018	0.024	0.031	0.029	0.040
	(0.011)	(0.010)	(0.020)	(0.020)	(0.020)	(0.024)	(0.011)	(0.011)	(0.022)	(0.023)	(0.022)	(0.029)
Male	0.117	0.107	0.096	0.169	0.110	0.088	0.205*	0.189*	0.065	0.056	−0.019	0.040
	(0.090)	(0.086)	(0.150)	(0.162)	(0.160)	(0.180)	(0.106)	(0.105)	(0.184)	(0.200)	(0.196)	(0.210)
Preschoool teachers	0.096	0.069	0.188	0.103	0.188	0.072	0.091	0.058	0.024	−0.026	0.014	−0.086
	(0.102)	(0.099)	(0.144)	(0.144)	(0.148)	(0.179)	(0.122)	(0.118)	(0.182)	(0.195)	(0.188)	(0.219)
Imm ackground	−0.030	−0.018	0.042	0.083	0.026	−0.049	0.013	0.024	0.106	0.101	0.083	0.036
	(0.071)	(0.072)	(0.113)	(0.120)	(0.112)	(0.137)	(0.078)	(0.078)	(0.133)	(0.142)	(0.134)	(0.149)
Sick-days	−0.180	−0.127	−0.439	−0.328	−0.395	−0.403	−0.016	0.039	−0.394	−0.309	−0.164	−0.255
	(0.172)	(0.167)	(0.279)	(0.296)	(0.293)	(0.316)	(0.183)	(0.180)	(0.309)	(0.312)	(0.320)	(0.363)
Ln(nr children)	−0.007	−0.007	−0.006	−0.013	−0.008	−0.018	−0.020	−0.019	−0.005	−0.015	−0.012	0.005
	(0.019)	(0.019)	(0.031)	(0.034)	(0.033)	(0.039)	(0.025)	(0.024)	(0.036)	(0.036)	(0.037)	(0.044)
Nr children/nr employee	−0.014	−0.011	−0.019	−0.020	−0.011	−0.012	−0.004	−0.001	−0.012	−0.016	−0.004	−0.002
	(0.011)	(0.011)	(0.018)	(0.018)	(0.018)	(0.020)	(0.012)	(0.011)	(0.019)	(0.021)	(0.019)	(0.023)
Distance					0.003						−0.004	
					(0.008)						(0.009)	
Controlling for:												
- Fam char	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
- Choice set	No	No	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes
- School FE	No	No	No	Yes	No	No	No	No	No	Yes	No	No
Keep those who accepted 1st offer						Yes						Yes
Nr of obs	2175	2175	2175		2121	1896	2175	2175	2175		2121	1896

Note: The table reports OLS estimates from estimating different variants of Eq. (1) for mathematics. All child care characteristics are measured at the child care center level. As a measure of tenure we use average years the current employees have been working in a given center. Other staff characteristics are measured as shares.
Included in all specifications are dummy variables for parents' preferences where we control separately for the first ranked center, a constant term and the family background variables listed in Section 3.1 and Appendix Table 10, (staff's) tenure squared and dummy variables for the child's birth year and month, mother's continent of origin, number of ranked centers, application year, city districts and interaction terms between application year and city districts. Standard errors are clustered at the (first offered) center level and are robust to heteroskedasticity. */**/*** denotes statistical significance at the 10/5/1% level.

References

Ahnert, L., Pinquart, M., Lamb, M.E., 2006. Security of children's relationships with nonparental care providers: a meta-analysis. Child Dev. 77 (3), 664–679.

Almond, D., Currie, J., 2011. Human capital development before age five. Part B. Elsevier. In: Card, D., Ashenfelter, O. (Eds.), Handbook of Labor Economics. 4, pp. 1315–1486.

Antecol, H., Eren, O., Ozbeklik, S., 2015. The effect of teacher gender on student achievement in primary school. J. Labor Econ. 33 (1), 63–89.

Araujo, M.C., Carneiro, P., Cruz-Aguayo, Y., Schady, N., 2016. Teacher quality and learning outcomes in kindergarten. Q. J. Econ. 131 (3), 1415–1453.

Baker, M., 2011, November. Innis lecture: universal early childhood interventions: what is the evidence base? Can. J. Econ. 44 (4), 1069–1105.

Bauchmuller, R., Goertz, M., Rasmussen, A.W., 2014. Long-run benefits from universal high-quality preschooling. Early Child. Res. Q. 29, 457–470.

Blau, D.M., 1999. The effect of child care characteristics on child development. J. Hum. Resour. 34 (4), 786–822.

Blau, D., Currie, J., 2006. Pre-school, day care, and after-school care: who's minding he kids? Volume 2 of Handbook of the Economics of Education, Chapter 20. Elsevier, pp. 1163–1278

Bowlby, J., 1969. Attachment and Loss: Vol.1. Attachment. Basic Books, New York.

Boyd-Swan, C., Herbst, C.M., 2017. Racial and ethnic discrimination iin the labor market for child care teachers. IZA Discussion Papers 11140. Institute for the Study of Labor (IZA).

Camilli, G., Vargas, S., Ryan, S., Barnett, W.S., 2010. Meta-analysis of the effects of early education interventions on cognitive and social development. Teach. Coll. Rec. 112 (3), 579–620.

Choi, J.Y., Horm, D., Jeon, S., Ryu, D., 2019. Do stability of care and teacher–child interaction quality predict child outcomes in early head start? Early Educ. Dev. 30 (3), 337–356.

Clotfelter, C.T., Ladd, H.F., Vigdor, J.L., 2009. Are teacher absences worth worrying about in the United States? Educ. Finan. Policy 4 (2), 115–149.

Currie, J., Neidell, M., 2007, February. Getting inside the "black box" of head start quality: what matters and what doesn't. Econ. Educ. Rev. 26 (1), 83–99.

Dee, T.S., 2004. Teachers, race, and student achievement in a randomized experiment. Rev. Econ. Stat. 86 (1), 195–210.

Dee, T.S., 2005. A teacher like me: does race, ethnicity, or gender matter? Am. Econ. Rev. Pap. Proc. 95, 158–165.

Dee, T.S., 2007. Teachers and the gender gaps in student achievement. J. Hum. Resour. XLII (3), 528–554.

Drange, N., Havnes, T., 2019. Early child care and cognitive development: evidence from an assignment lottery. J. Labor Econ. 37 (2).

Drange, N., Telle, K., 2018. Child care and inequality of opportunity. Descriptive findings from Norway. Discussion Papers 879. Research Department of Statistics Norway.

Fletcher, R., StGeorge, J., Freeman, E., 2013. Rough and tumble play quality: theoretical foundations for a new measure of father–child interaction. Early Child Dev. Care 183 (6), 746–759.

Goertz, M., Johansen, E.R., Simonsen, M., 2018. Academic achievement and the gender composition of preschool staff. Labour Econ. 55, 241–258.
Hoffmann, F., Oreopoulos, P., 2009. A professor like me: the influence of instructor gender on college achievement. J. Hum. Resour. 44 (2), 479–494.
Horm, D.M., File, N., Bryant, D., Burchinal, M., Raikes, H., Forestieri, N., Encinger, A., Cobo-Lewis, A., 2018. Associations between continuity of care in infant-toddler classrooms and child outcomes. Early Child. Res. Q. 42, 105–118.
Knudsen, E.I., Heckman, J.J., Cameron, J.L., Shonkoff, J.P., 2006. Economic, neurobiological, and behavioral perspectives on building America's future workforce. Proc. Natl. Acad. Sci. U. S. A. 103 (27), 10155–10162.
Krueger, A., 1999. Experimental estimates of education production functions. Q. J. Econ. 114 (2), 497–532.
Norwegian Ministry of Education, 2006. Rammeplan for Barnehagens Innhold Og Oppgaver (An Overall Plan for the Content and Activities in the Child Care Center). Norwegian Ministry of Education.
OECD (2017). Starting Strong 2017: Key OECD Indicators on Early Childhood Education and Care. OECD Publishing. of Occupational Health, N. I. (2016). Statistics from the department of occupational health surveillance (noa). Technical report.
Pellegrini, A.D., Smith, P.K., 1998. Physical activity play: the nature and function of a neglected aspect of play. Child Dev. 69 (3), 577–598.
Rivkin, S.G., Hanushek, E.A., Kain, J.F., 2005. Teachers, schools, and academic achievement. Econometrica 73 (2), 417–458.
Ruhm, C.J., Waldfogel, J., 2012. Long-term effects of early childhood care and education. Nordic Econ. Pol. Rev. 1, 23–51.
Sroufe, L.A., Coffino, B., Carlson, E.A., 2010. Conceptualizing the role of early experience: lessons from the Minnesota longitudinal study. Dev. Rev. 30 (1), 36–51.
Storli, R., Sandseter, E.B.H., 2017. Gender matters: male and female ECEC practitioners' perceptions and practices regarding children's rough-and-tumble play (R and T). Eur. Early Child. Educ. Res. J. 25 (6), 838–853.
Walters, C.R., 2015. Inputs in the production of early childhood human capital: evidence from head start. Am. Econ. J. Appl. Econ. 7 (4), 76–102.