



Boys left behind: The effects of summer camp and follow-up strategies on academic, personal, and social competencies

Søren Albeck Nielsen*

TrygFonden's Centre for Child Research, Aarhus University, Denmark
 Department of Economics and Business Economics, Aarhus University, Denmark
 Centre for Integrated Register-based Research, CIRRAU, Aarhus University, Denmark

ARTICLE INFO

JEL classification:

I21
 I24
 I28
 C23

Keywords:

Educational economics
 Intensive learning camps
 Follow-up strategy
 Mentoring
 Non-cognitive competencies

ABSTRACT

I use Danish administrative data to investigate the consequences of summer camp participation combined with a one-year follow-up program for disadvantaged boys on academic, personal, and social competencies. My identification strategy relies on individual-level panel data that enable me to observe outcomes before and after summer camp participation. Using a difference-in-differences strategy, I find overall positive effects on academic and personal competencies that reduce the gap to a matched group of boys with similar background characteristics by 40 to 80 percent. Further, I exploit a structural change in the follow-up program to evaluate how different mentoring strategies affect outcomes. In 2017, the follow-up program was changed from individual mentoring to group mentoring. Using a triple differences strategy, I find that group mentoring in the follow-up program improves personal and social competencies, suggesting that the format of the follow-up program is crucial for effects on personal and social competencies.

1. Introduction

This paper studies the consequences of a treatment package consisting of summer camp and follow-up program participation for disadvantaged boys in lower secondary education. Furthermore, I study two different structures of the follow-up program. In recent decades, there has been a rapid and substantial reversal of the gender gap in educational attainment in much of the developed world (Murnane, 2013). The OECD (2013) report shows that, in 28 of 34 OECD countries, females have surpassed males in higher education among adults aged 25 to 34. Autor et al. (2019) conclude that boys from disadvantaged households have higher rates of disciplinary problems, lower academic performance, and lower high school completion rates compared to girls with similar backgrounds. Additionally, boys from low socioeconomic backgrounds have been particularly affected by the disruption to education caused by the Covid-19 pandemic (Di Pietro et al., 2020). Thus, there is a need for effective remedial education programs to eliminate this learning gap for disadvantaged boys.

Intensive learning programs are widely used as a policy tool to increase educational attainment and are often implemented as summer camps (Cooper et al., 2000; Kim & Quinn, 2013; Lauer et al., 2006). The three meta-analyses report positive short-run effects on math and reading, but with very small effect sizes. Despite the overwhelming

use of summer camps in society, there is little causal evidence. Especially, on the combination of summer camp and a follow-up program. Only a handful of studies rely on experiments or quasi-experiments, often with low statistical power, i.e. total sample size below 100. The most convincing evidence on summer camps with large sample sizes exploits regression discontinuity designs (henceforth RDD) and standardized tests using the threat of grade retention, but do not have a follow-up component. Mariano and Martorell (2013) find modest effects on language and little effect on math for 4,000 low-performing pupils of which 38% belong in the treatment group. The RDD of Jacob and Lefgren (2004) exploits the entire population of pupils in Chicago public schools and finds positive effects on both reading and math, but only for grade 3 pupils and not grade 6 pupils. Recent evidence from Battistin and Schizzerotto (2019) finds negative effects on academic performance of mandatory summer camp for at-risk pupils in Italy using a sample of around 1,800 pupils with half receiving treatment. Overall, the evidence on the effect of summer camps on boys during the transition from lower to upper secondary education is weak, and there is no evidence on the effect of summer camps on non-academic outcomes in general as well as in combination with a follow-up program.

* Correspondence to: Aarhus University, Fulgesangs Alle 4, 8240 Aarhus V, Denmark.

E-mail address: snielsen@econ.au.dk.

¹ Indicator variable taking the value one if ready and zero otherwise. See Section 3 for the description of the readiness assessment.

I use population-level Danish register data covering all grade 8 boys in the period 2015–2019 to study the effect of a two-week summer camp for academically disadvantaged boys with a one-year follow-up mentoring program. In particular, I exploit grade 8 and 9 individual-level panel data to implement a difference-in-differences strategy. I ask, first, what are the effects of camp/mentoring participation on academic performance and readiness¹ for upper secondary education? Secondly, I provide novel evidence on the consequences of changing the follow-up program from individual mentoring with an adult to a group-mentoring program of 10 to 12 boys.

The analyses deliver a set of noteworthy answers to the research questions: First, I show that participating in a two-week summer camp with a one-year follow-up during the holiday between grades 8 and 9 has positive impacts on academic performance and readiness for upper secondary education. I observe an increase in the average assessment mark of around 15% of a standard deviation, suggesting that overall academic abilities have improved on average. Additionally, participation in the camp/mentoring increases readiness for upper secondary education by 18 percentage points at the educational readiness assessment in December grade 9. The overall readiness assessment consists of three sub-categories in which the pupils must qualify in order to be ready for upper secondary education. Interestingly, academic readiness is the main reason for the overall effect with an increase of 22 percentage points, whereas the summer camp increases the personal readiness assessment by 10 percentage points, and the effect on social readiness is not statistically significant. This estimation strategy combined with the introduction of a group-mentoring follow-up program in 2017 allows me to identify the effect of moving from individual mentoring with one adult to a group-mentoring follow-up program if nothing else changes. Using a triple differences strategy, I find that substituting individual mentoring with group mentoring increases both personal and social readiness assessments by 15 and 16 percentage points, respectively, and has no impact on academic readiness. This indicates that participating boys benefit both personally and socially from staying connected with equal peers and sharing experiences and solutions. Intuitively, it makes sense that mentoring these boys together should benefit their behavior, but one might also fear that connecting challenged boys on a regular basis might increase poor behavior. [Petrosino et al. \(2013\)](#) show exactly this. They find that crime preventive camps for children at risk of becoming delinquent do not work and in fact increase poor behavior.

Thus, my analysis shows that at-risk boys improve their academic, personal, and social competencies through participation in a two-week intensive learning camp during the summer holiday between grades 8 and 9 with a one-year follow-up program. These larger results compared to the previous literature on summer camps might, at least, in part be explained by the combination of summer camp and a one-year follow-up program. Importantly, I find that the improvement in non-academic competencies is a result of shifting the follow-up program from individual mentoring to group mentoring. This finding is particularly important as it indicates that intensive learning programs aimed at at-risk boys should focus on group mentoring in the follow-up program to boost the benefits of the camp/mentoring package.

I contribute to the literature on the consequences of intensive learning camps. Firstly, often the effect is measured shortly after the camp is completed. In this study, however, I exploit administrative data to investigate the effect up to one year after camp participation. Secondly, the educational readiness assessment allows me to supplement the estimated effects on academic outcomes with effects on measures of personal and social competencies directly related to future educational attainment six months after the camp. These outcomes are under-investigated in the previous literature, which is problematic when considering the importance of non-academic skills for future achievements ([Almlund et al., 2011](#)). Finally, the change in the follow-up strategy allows me to add novel evidence on how follow-up programs should be structured when combined with intensive learning camps with the attempt to maximize the effect on all competencies that are

considered relevant for future educational outcomes — and not focus exclusively on the academic competencies.

The paper is organized as follows: Section 2 provides background information on the structure of the summer camp. Section 3 describes the data and descriptive statistics and Section 4 explains the empirical strategy. In Section 5, I show the effect of the summer camp, Section 6 investigates the effect of changing the follow-up strategy, and finally, Section 7 concludes.

2. The two weeks summer camp

The treatment program is a two-week intensive learning camp that takes place during the summer holiday between grades 8 and 9 with a one-year follow-up program consisting of mentoring. It targets boys who are at risk of not becoming ready for upper secondary education at the end of compulsory schooling and, thus, not able to choose their desired education. This group comprises boys with a substantial academic backlog, low self-confidence, and low motivation for going to school. The main objective of camp/mentoring is to help these academically weak pupils to catch up such that they become prepared to continue in the education system. In order to achieve this objective, the program focuses on academic abilities by training reading, writing, spelling, and math, but it also aims to develop socio-emotional skills.

2.1. Content and organization

The recruitment into the program occurs when the boy in collaboration with their parent apply for the camp independently from their local school. Applications can be submitted from January, with a deadline of mid-March. The boys are notified of acceptance within a month from the deadline and all applications are evaluated simultaneously. If the program is over-subscribed the boys are selected based on academic, personal and social problems. The boys must not be in active problems with the legal system.

The summer camp consists of an introductory camp and a main camp. About two weeks before the start of the main camp, pupils attend an introductory camp over a weekend. The purpose of the intro camp is to introduce the teachers, structure, settings, and rules at the main camp as well as create relationships with their peers in order to start the social connections and decrease some of the anxiety before the main camp. Additionally, there are individual conversations between the teachers and the boys for the teachers to get a greater insight into the boys' difficulties and potential. Finally, there are academic tests of abilities in order to organize the main camp's learning plan according to the pupils' individual levels and needs ([Andersen et al., 2019](#)). The total cost for the summer camp, including the follow-up program, is approximately 5,000 USD per pupil.

The main camp takes place in the first two weeks of the summer holidays. The boys are divided into teams of approximately 15 pupils, where two teams receive instruction together. The two-week camp consists of a full schedule from morning to evening with a morning wake-up call at 6.30 AM and bedtime at 10.00 PM. This ensures a fixed structure that is repeated every day. The day consists of four 90-minute modules, which include the subjects Danish, and mathematics, as well as modules focusing on socio-emotional skills and plenty of physical activity. See Figure A.1 for an example of the schedule for the first week of the main camp. The camp is located at a boarding school² with teaching facilities, a kitchen, and sleeping rooms. Thus, there is no need to leave the camp during the two weeks. The teachers and peers are different from the participant's local learning environment, mobile phones are only allowed one hour per day, and the camp has a zero-sugar policy. Importantly, the boys do not miss any teaching at

² Same boarding school in all years.

their home school, because the intro camp is held during a weekend and the main camp during the summer holiday.

The pedagogical aim of the summer camp is to offer boys teaching methods that, to a greater extent than at their home school, are tailored to the boys' individual needs and strengths. The intervention builds on inspiration from the literature on Visible Learning and Self-Determination Theory (Hattie & Yates, 2013; Ryan & Deci, 2000). The teaching itself is highly structured around Flipped classrooms, where blackboard teaching is replaced by e.g. short videos. The boys watch the videos individually to free up the teacher's time for more student-activating teaching. An additional advantage of this method is that this form of teaching also means that the boys can revisit the videos etc. after the camp. The teachers at the camp attempt to turn around the negative school experiences by testing the boys midway through the camp and on the final day. Thus, they visually illustrate the boys' academic progress to recreate a positive self-narrative as well as the motivation to learn. A clear pedagogical tool of the learning camp is to believe in the boys by praising and acknowledging them for their progress and supporting them during difficult periods.

2.2. The follow-up strategy

Receipt of educational activities in a new environment, away from the home-school peers and teachers, is a key element of the program. The intention is to break bad habits and make new social connections, exploiting that all boys are on common ground by not knowing each other beforehand. This is an obvious strength of the learning program, but at the same time, it also constitutes one of the biggest challenges. Andersen et al. (2019) show that many pupils find it difficult to maintain good working habits, positive academic development, and the joy of learning when they return to their ordinary classroom after the camp. The summer camp is seen as being disconnected from the daily school environment, and thus, returning to their local surroundings could result in a setback if the boys find that their peers and local teachers still view them in the same manner. Therefore, the program has a great focus on follow-up after the camp. The primary follow-up program is a mandatory one-year mentoring scheme with a built-in parenting effort. Additionally, the camp teachers construct an "Exit package" for all boys, which is a written handover to the boys' local schoolteachers. However, only 50% of the local teachers have heard about the exit package and fewer have used it (Rambøll, 2019). There is no additional formal contact between the camp teachers and the local teachers. In 2015 and 2016, the mentoring scheme was an individual mentor-mentee program, where the boy and the adult mentor can train in academic as well as non-academic skills during meetings twice a month. Additionally, the mentors can help the boys set new learning goals and maintain their positive development. In 2017, the mentoring scheme was changed to mentor groups of 10–12 boys, who meet twice a month, outside regular school hours, at mentor centers geographically spread across Denmark. At the centers, the boys receive academic and personal counseling and guidance from adults who have attended the summer camp as well as from volunteer role models. Furthermore, the content of the summer camp is repeated at each meeting, they share experiences related to returning to the local learning environment, and they receive homework help and educational guidance.

3. Data and descriptive statistics

To investigate the effect of the program package and how changes in the follow-up strategies affect the outcomes of participants, I leverage Danish administrative register data available through Statistics Denmark covering the full population of pupils in the Danish school system. Focal to this study is the Danish Student Register comprising all educational choices in Denmark. This register is a unique longitudinal dataset that allows me to follow schooling information such as private vs public schooling, school and classroom movements, and special

needs teaching from 2008/2009 to 2019/2020. Crucially, this data is informative about what grade a boy attends, enabling me to observe the boys' academic development through their average assessment mark and their readiness assessment for upper secondary education in grades 8 and 9 obtained from the Ministry of Education. I augment this data with information on dyslexia, psychiatric diagnoses, results from national tests in reading and math, school absence, socio-emotional skills, and school well-being. Furthermore, I exploit socioeconomic information describing the demographics, employment, income and educational level of the parents. The sample consists of 158,231 boys of which 241 participated at the summer camp between 2015 and 2019. I investigate if the large differences in sample size and composition between the camp boys and the comparison boys affect the results by implementing a different reweighing strategy that equalizes the two groups. The conclusions are robust to all approaches (See Section 5). The 241 boys are those who applied, were invited, and attended the main summer camp. Due to legal issues, the camp provider only registered boys who attended the main camp. Thus, I do not know who was invited and who did not show up on the first day. Starting in 2017, the program began recording dropouts from the camp. Approximately, 8% left during the two weeks of camp either due to bad behavior or because they did not want to be there. Thus, I am able to estimate the intention-to-treat effect for those who showed up at the first day of camp.

3.1. Average assessment mark

The average assessment mark is an expression of the pupil's academic level averaged across all subjects the pupil attends.³ The teachers evaluate the student in relation to the subject's academic goals. They base their assessment on the degree to which the student has achieved the subject's competence, skill, and knowledge objectives. Importantly, the pupil's work effort and/or behavior in the classroom is, generally, not included in the grading. Only if the subject has goals that include these abilities, are they included in the grading. The grading occurs in December and June of grades 8 and 9.⁴

To measure the pupils' overall abilities, I standardize each continuous assessment mark within each subject and the timing of the assessment to mean zero and a standard deviation of one. Then, I calculate the average assessment mark and standardize it within the timing of the assessment to mean zero and a standard deviation of one. The latter standardization allows me to readily interpret regression coefficients in standard deviation units and, thus, render results comparable to the effect sizes of other studies. The camp boys are on average 1.2 standard deviations below the non-camp boys at both grade 8 assessments, indicating that camp boys have a significant academic backlog.

3.2. Readiness assessment for upper secondary education

Assessing pupils' readiness in relation to choosing and completing upper secondary education is a process that starts in grade 8. The purpose of the assessment is to ensure that non-ready pupils receive school interventions and individual guidance in the process towards the end of grade 9 in order for them to make the best secondary education choice and be prepared for the education chosen. The readiness assessment

³ In Danish, there are four grades (Reading, Spelling, Written, and Oral). In math, there are three grades (Math with aids, Math without aids, and Oral). In foreign languages, there are two grades (Oral and Written). For the following mandatory subjects, there is only one grade (Physics/chemistry, Biology, Geography, History, Social studies, Religion, and Gym). Finally, the pupils are graded in one of the following four electives (Crafts and design, Food Knowledge, Music, and Art).

⁴ The June grade 8 grading is initiated in 2018 and does not exist for the previous years.

includes all pupils in public and private schooling and takes place only in December. The readiness assessment was first conducted in 2015 and a reform of the readiness assessment in 2019 allows only for the use of this four-year period. The pupil's primary teachers assess the academic, personal, and social competencies of the pupil. All three criteria must be met for the pupils to be assessed as ready for upper secondary education. On average, 62% of the boys are assessed to be ready for upper secondary education, but for the camp participants in grade 8, only 12% are assessed to be ready for upper secondary education. This 50%-point difference clearly indicates that the camp participants need additional assistance in order to continue in the educational system.

In order to be academically ready in grade 8, the pupil must have an average assessment mark of at least 4.0⁵ for going to vocational training and two-year high school and 5.0 for a three-year high school. In grade 9, the requirements for vocational training drop to an average of 2.0 in Danish and math. In grade 8, 75% are academically ready for their desired educational choice, whereas this fraction is only 19% for the camp participants.

For the personal readiness assessment, the teachers assess whether the pupil has the necessary personal competencies to begin upper secondary education after grade 9. The teachers have five key areas to guide them through the assessment: (1) Motivation, (2) Independence, (3) Responsibility, (4) Meeting stability, and (5) Choice readiness. Motivation refers to the eagerness for education, a passion for learning, and active participation in the teaching process. Independence is about being able to act on your own, i.e. the pupil is able to take initiative and ask for help when needed. Responsibility is whether the pupil shows up prepared for classes and whether the pupil is able to keep appointments that he or she makes with teachers or fellow pupils. Meeting stability is whether the pupil attends school every day and does so timely. Finally, choice readiness is about whether the pupil can make decisions and whether the pupil is able to make a positive and active choice in the educational selection process. Often a pupil will not be able to make the educational choice in grade 8, but they must be able to reflect on it. In grade 8, 74% of all boys are assessed as personally ready, while the fraction is 40% for the camp participants. The camp boys still lag behind the average boy, but the gap is not as significant as in the academic readiness assessment.

The social readiness assessment focuses on whether the pupil has the social prerequisites needed to be able to start and complete upper secondary education. In order to assess this, the teachers have three focus points to guide them: (1) Collaboration ability, (2) Respect, and (3) Tolerance. Collaboration ability is about being able to solve tasks together with others, keep common agreements, and contribute positively to the community. Respect is about the pupil being able to show consideration for other pupils and teachers. Tolerance is the ability or willingness to accept what is unknown. Thus, tolerance is about being able to understand and accept other people's opinions, behavior, culture, religion, etc. On average, 82% of all boys are assessed to be socially ready, while 56% of the camp boys are socially ready.

It is evident that the significant difference in the overall readiness assessment is primarily due to poor academic performance. A considerable portion of the camp participants also face personal and social issues, but not to the same extent as their academic deficits.

3.3. Characteristics of camp vs non-camp boys

Leveraging Danish administrative data allows me to characterize in detail the differences between camp and non-camp boys. In columns 1 to 5 of Table A.1, I present summary statistics for a rich set of

covariates, recorded prior to camp participation. The descriptive statistics show significant differences between the two groups. There is an under-representation of boys with non-western ethnicity in the camp group. They are more likely to live in broken families, have additional relocations, and thus have an increased number of school changes. The camp boys have received on average half a year more special needs teaching, they are over-represented in terms of psychiatric diagnoses, and 35% of them have been diagnosed with dyslexia, whereas only 10% of the non-camp group are dyslexics. They grow up in households with lower socioeconomic status, with both parents having approximately half a year less of education. Academically, the camp group performs significantly worse in the national test in reading and math across all grades and profile areas with gaps to non-camp boys between 0.5 and 1.0 standard derivations. The camp participants on average rate their school well-being lower, especially when assessing their learning self-efficacy. I find similar results for all three evaluated socio-emotional skills, with the participants who scored lower on conscientiousness exhibiting the largest gap of 0.5 standard deviations.

4. Empirical strategy

4.1. Identifying the consequences of program participation

The first goal of this paper is to estimate the consequences of summer camp and mentoring participation for all camp boys⁶ on pupil-level outcomes. The key challenge in any program evaluation is to estimate the counterfactual, in the present case the outcomes in the absence of camp/mentoring participation. A natural worry is that selection bias challenges the identification, i.e. camp/mentoring participating boys comprise a different population compared to the remaining population of boys who do not participate in the camp/mentoring program.

I address this concern with a difference-in-differences strategy using individual-level panel data similar to an individual fixed effect analysis. This strategy compares the change in participating pupils' outcomes from grade 8 to grade 9 to a similar change for non-participating pupils. This strategy implicitly controls for unobserved time-invariant individual school performance. However, it is likely that school performance develops differently based on the underlying distribution of covariates. To account for this, I combine the difference-in-differences strategy with Entropy Balancing along the lines of Freier et al. (2015). The basic idea is to reduce bias due to different distributions of covariates in the camp/mentoring vs non-camp/mentoring groups by matching the non-camp/mentoring group such that it is identical to the camp/mentoring group with respect to included characteristics (Abadie, 2005; Blundell et al., 2004; Heckman et al., 1997). Entropy Balancing is a data processing method introduced by Hainmueller (2012) that ensures perfect covariate balance with a binary treatment variable. See Hainmueller and Xu (2013) and Hainmueller (2012) for a detailed description of Entropy Balancing. Columns 6 to 8 in table A.1 show descriptive statistics of the observed covariates for the Entropy Balancing adjusted non-camp/mentoring group. Column 8 shows that the means in the camp/mentoring and non-camp/mentoring groups are perfectly balanced across the full set of covariates.

I start the analyses with the following difference-in-differences equation:

$$y_{it} = \alpha_i + \beta_0 + \beta_1 \text{Camp}_i + \beta_2 \text{Time}_i + \delta_j (\text{Camp}_i \cdot \text{Time}_i) + \epsilon_{it} \quad (1)$$

where y_{it} is the outcome of interest, Camp_i is a binary variable indicating camp/mentoring participation, and Time_i is a time-period categorical variable containing two periods before and after camp/mentoring

⁵ Academic grading in Denmark is on a 7-point scale with the following grades from the best to the worst: 12(A), 10(B), 7(C), 4(D), 02(E), 00(FX), -3(F).

⁶ I pool together all grade 8 boys in the school years 2014/2015 to 2018/2019 when estimating the overall consequences of summer camp participation.

participation for the average assessment mark and one for the readiness for upper secondary education assessment. δ_j are the effects of camp/mentoring participation by time periods relative to December in grade 8 and are the parameters of interest, i.e. the intention to treat (ITT henceforth) of those showing up at the first day of the summer camp. u_i is individual level fixed effect and ϵ_{it} is the error term. Standard errors are clustered at the individual level as outlined in Bertrand et al. (2004).

The key identifying assumption in the difference-in-differences strategy is that there can be no differential trends between the camp/mentoring and non-camp/mentoring groups in the absence of program participation. To investigate the validity of this assumption, I first observe pre-program trends for the average assessment mark and then test the robustness of my findings to different model specifications, alternative comparison groups, alternative balancing specifications, and by performing placebo analyses prior to camp participation and on a randomly selected, synthetic, “camp/mentoring” group. Significant and large placebo effects would imply that the identification strategy does not capture systematically different trends in pre-program school performance between camp/mentoring and non-camp/mentoring boys.

4.2. Detecting the consequences of transforming the follow-up program

The second goal of this paper is to estimate the impact of transforming the follow-up program – the 2017 structural change in the mentoring scheme – on pupil-level outcomes. Individual mentoring is substituted with group mentoring in the one-year mandatory follow-up program. The effect of this change can be estimated using a triple difference estimator.⁷ This is equivalent to the difference between 2015 to 2016 and 2017 to 2019 difference-in-differences estimates and is estimated using the following equation:

$$y_i = \beta_0 + \beta_1 \text{Camp}_i + \beta_2 \text{Post}_i + \beta_3 \text{Group}_i + \beta_4 \text{Camp}_i \cdot \text{Post}_i + \beta_5 \text{Camp}_i \cdot \text{Group}_i + \beta_6 \text{Post}_i \cdot \text{Group}_i + \beta_7 \text{Camp}_i \cdot \text{Post}_i \cdot \text{Group}_i + \beta_8 X_i + \epsilon_i \quad (2)$$

where y_i is the outcome of interest, camp_i is a binary variable indicating camp/mentoring participation, Post_i is an indicator variable with the value one (zero) after (before) camp/mentoring participation, Group_i takes the value one for boys participating in the group mentoring scheme and zero for boys in the individual mentoring scheme, X_i is a matrix containing covariates measured prior to program participation, and ϵ_i is the error term. β_7 is the effect of camp participation under the group mentoring follow-up program relative to individual mentoring and is the parameter of interest.

Despite that the triple differences estimator can be computed as the difference between two difference-in-differences estimators, Gruber (1994) states that the identifying assumptions are weaker. Olden and Møen (2020) formally shows this by proving that the triple differences estimator does not require two parallel trend assumptions in order to estimate causal effects. It requires only one parallel trend assumption to hold for causal interpretation. Thus, the difference between two biased difference-in-differences estimators will not be biased if the bias is the same in both estimators because the bias will be removed with triple differencing.

Even though the 2017 reform of the follow-up program provides a unique possibility to evaluate two treatment packages where both include the same summer camp, but with two different follow-up programs, some assumptions need to be met for causal interpretation. To establish a credible belief that the difference in the impact between the two intervention packages can indicate which type of mentoring, individual or group, is more effective, it is necessary for the summer

camp to maintain a similar level of quality before and after the 2017 reform. All summer camps are held at the same boarding school location and the structure of the course has not changed. There are on average 23.7 teachers at the camp before the reform and 23.5 after the reform and the average number of new teachers for each camp are 8 before the reform and 9 after the reform indicating that the teaching quality of the two weeks summer camp is similar across the reform. Another assumption is that the program has not evolved over time. This is challenging to verify, but one could anticipate that the enhancement of a camp/mentoring program like this would level off with time. The summer camp under examination started in 2012 and has been in operation annually, so it is not probable to anticipate significant improvement in the program from 2015 to 2019 due to the accumulation of experience.⁸ Finally, one might worry that the selection process changes with the reform. This is, however, not the case (see table A.15). Even so, I cannot completely rule out other things correlated with the change in the mentoring program and the triple difference estimates must be interpreted with caution.

5. Results: Effect of camp/mentoring participation on school performance outcomes

I start with a graphical analysis of the effects on the average assessment mark measured four times across grades 8 and 9. Here I use camp/mentoring participation for all the years 2015–2019. Fig. 1 shows the event study representation of camp/mentoring participation on the average assessment mark with December grade 8 as the reference point. The solid line presents the unadjusted event study (raw difference-in-differences estimates) and the dashed line presents the entropy balancing adjusted event study. Importantly, there is no difference in effect between December and June in grade 8, aligning with no violation of the parallel trend assumption. Additionally, the adjusted and unadjusted treatment effects are very similar across time, indicating that the different distribution of covariates does not affect the estimated effect of camp participation. Overall, I find effects of 15% of a standard deviation both 6 and 12 months after camp/mentoring participation, suggesting that the effects on academic grades are persistent up to a year after the intervention. Participating in the two-week summer camp with a one-year follow-up reduces the learning gap to the full population of boys by 14% and by 40% when compared to boys with similar characteristics.

A worry might be that schoolteachers who are grading the boys in the average assessment mark might be influenced by the knowledge of who has participated in the program. Therefore, I exploit data from the grade 9 school leaving exams in June that are graded by external sensors to investigate if this changes the results. Figure A.3 shows the Entropy Balancing adjusted event study representation, where the average assessment mark in June of grade 9 is substituted with the exam performance. This worry does not seem to be valid since using exam performance does not change the findings. To investigate this further, I use the written exams because these are completely independent of the teachers. Figure A.4 supports the conclusion that the teachers' beliefs about the camp/mentoring program do not influence the effects.

For simplicity and similarity in presentation with the readiness assessment for upper secondary education, which is only registered once in grades 8 and 9, I supplement my event study model with a simple difference-in-differences model that ignores the time to treatment aspect, i.e. it does not allow effects to vary with time distance to summer camp. I also conduct robustness checks using this simpler difference-in-differences model. This is not critical for my findings, conducting the robustness checks on the main event study model yields similar findings. Table A.2 shows in columns 1 to 3 the unadjusted

⁷ Individual fixed effect estimates are no longer possible because none of the boys has participated more than once in the program.

⁸ The program started in 2012, but due to data limitation in respect to outcomes I am only able to use the 2015 to 2019 cohort.

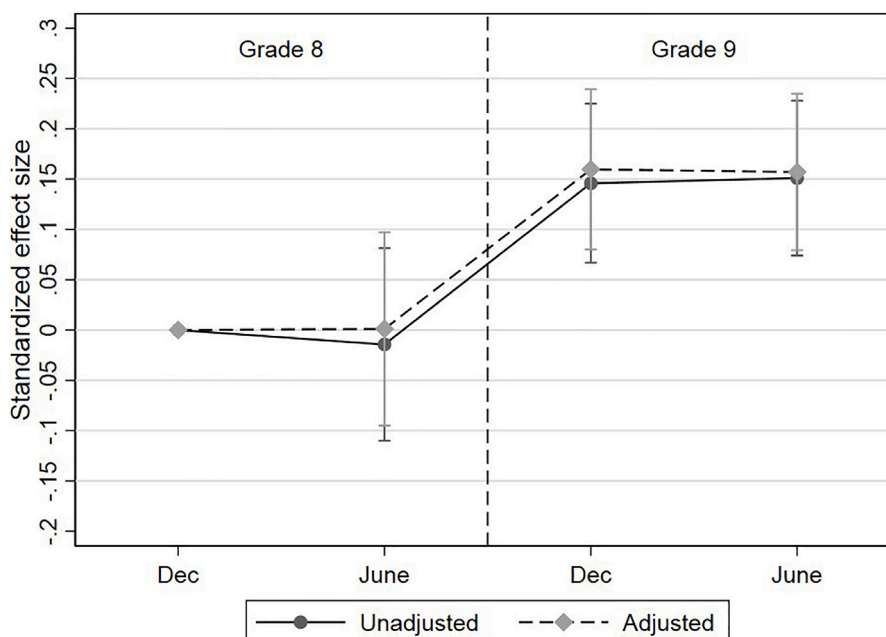


Fig. 1. Camp/mentoring and average assessment mark.

Notes: This figure shows the main event study representation of the effect of camp/mentoring participation on average assessment mark. The solid line presents the unadjusted event study and the dashed line presents the Entropy Balancing adjusted event study. Each point represents the assessment difference in average assessment mark between camp/mentoring and non-camp/mentoring boys with 95% confidence intervals. I use the assessment in December grade 8 as the reference time-point.

difference-in-differences estimations and in columns 4 to 6 the Entropy Balancing adjusted estimations. I present the raw estimations without covariates in Columns 1 and 4, in columns 2 and 5 I include covariates, and in columns 3 and 6 I exploit the panel structure of the data and conduct an individual fixed effects analysis. The table shows that the findings are robust to using a gradually richer specification.

Although the estimates are statistically significant it is important from a policy perspective to discuss if they are economically significant. Kraft (2020) discusses how to interpret effect sizes of educational interventions using 750 randomized trials. He argues that effects of 15% of a standard deviation are of medium size when benchmarked against a large set of educational interventions conducted in social science. However, when taking into consideration the age of the boys in this study, the effect on average assessment mark is around the 70 percentile of the distribution of effect sizes included in Kraft (2020). In a Danish context, this summer camp is also very competitive when compared to other lower secondary school interventions (Rosholm et al., 2021).

Table 1 explores the impact of camp/mentoring participation on the readiness assessment for upper secondary education. The results show that participation in the program leads to a considerable increase in readiness for desired upper secondary education. Participating boys increase their readiness by 18 percentage points relative to non-camp/mentoring boys. Remarkably, the camp/mentoring program reduces the gap to the full population of boys by 35% and by 80% to the entropy balance adjusted boys with similar characteristics. Columns 1 to 3 and 4 to 6 show that the findings are robust to using a gradually richer specification and combination with entropy balancing.

This result is consistent with the increase in the average assessment mark, which begs the question if the increased readiness for upper secondary education is purely driven by an increase in academic abilities. Tables A.3 to A.5 show the effect of camp/mentoring participation on the three criteria used in the readiness assessment. Clearly, the program has the largest effect on academic competencies with an increase of 22 percentage points. For personal competencies, the effect is borderline significant and shows an increase of 11 percentage points in the teacher-assessed personal readiness. Finally, the effect on social competencies is positive, however, not statistically significant.

However, as described in Section 3, the gap prior to the program is also the largest for the academic competencies. Thus, there is more room for improvement and maybe a larger focus on academic problems in the allocation of slots for the camp/mentoring program.

An important worry for the educational readiness assessment is that camp boys are able to change their desired upper secondary education between grades 8 and 9 and that this may cause positive effects. I.e. if the boys change from a 3-year high school to a vocational school, the requirements for being assessed ready are reduced. This does not seem to be a problem, as illustrated in table A.6; there are no effects on changing to another or easier upper secondary education and for the desired types of upper secondary education.

I also investigate the robustness of my findings to alternative comparison groups and balancing strategies. Using alternative comparison groups' works as further tests of the difference-in-differences combined with Entropy Balancing. I exploit the fact that the boys must apply for camp/mentoring participation with their parents and this is potentially related to local school interventions. Thus, if the local school does not have sufficient programs for the boys, they and their parents might be more inclined to apply for the program in order to adjust to their local schools' limitations. Therefore, I test the sensitivity of the results by using only boys from schools that have camp participants enrolled, as well as schools that have never sent pupils to the summer camp. The estimates in table A.7 indicate that the findings do not change. Comparing camp/mentoring participants to boys from their local schools or from different schools yields estimates with the same sign and magnitude as in the main specification. Additionally, I perform robustness analyses of the re-weighting approach in table A.8 to investigate if the findings are robust to alternative balancing methods. The first column of the table presents results from the main specification and balance on the first moment, column 2 balances on the first and second moments (mean and variance), and column 3 balances on three moments (mean, variance, and skewness). Columns 4 to 6 perform propensity score matching using logistic regression with 3, 5, and 10 non-camp/mentoring neighbor boys, respectively. Overall, the estimates have similar magnitudes and signs as the main specification, indicating that the analysis is robust to alternative balancing methods.

Table 1
Effects of summer camp on overall readiness assessment.

	Difference-in-difference			DiD & Entropy balance		
	(1)	(2)	(3)	(4)	(5)	(6)
Camp	-0.504 (0.022)	-0.251 (0.023)		-0.220 (0.022)	-0.221 (0.019)	
Post	0.147 (0.001)	0.139 (0.001)	0.139 (0.002)	0.206 (0.004)	0.190 (0.005)	0.193 (0.006)
Camp X Post	0.223 (0.038)	0.214 (0.038)	0.230 (0.060)	0.165 (0.038)	0.172 (0.037)	0.176 (0.060)
Observations	277,043	277,043	277,043	277,043	277,043	277,043
R-squared	0.026	0.352	0.820	0.119	0.359	0.763
Mean outcome, grade 9	0.776	0.776	0.776	0.522	0.522	0.522
Pupil background chars	No	Yes	-	No	Yes	-
Parental background chars	No	Yes	-	No	Yes	-
Reading abilities	No	Yes	-	No	Yes	-
Math abilities	No	Yes	-	No	Yes	-
Absence information	No	Yes	-	No	Yes	-
School well-being	No	Yes	-	No	Yes	-
Personality traits	No	Yes	-	No	Yes	-
Individual FE	No	No	Yes	No	No	Yes

Notes: This table presents the results for six separate difference-in-difference specifications comparing camp participants to non-participants. Columns 1 to 3 display the simple difference-in-difference specifications, and columns 4 to 6 show the results from the difference-in-difference combined with entropy balancing. Standard errors are clustered on the individual level. Missing values are imputed with the value zero and a binary indicator is added to the conditioning set. Bold (italic) numbers indicate significance at the 5% (10%) level.

Finally, I investigate the parallel trends assumption using placebo tests. First, I perform a placebo test using a fake treatment group on the main outcomes. The fake treatment group is a random subset of boys not affected by the program. Thus, estimates different from zero indicates a violation of the identification strategy. Table A.9 shows that all estimates are essentially zero. Another falsification test implements a fake camp date in the summer holiday between grades 5 and 6 for the boys participating in the camp between grades 8 and 9. To do so, I exploit the national tests in reading and math as additional outcomes. This fake camp date occurs prior to the real summer camp, and therefore the difference-in-differences strategy should find insignificant estimates close to zero in order to reject any underlying difference in trends between the camp and non-camp groups. Table A.10 shows small and insignificant estimates using national tests in reading, grades 6 and 8, and math grade 6, indicating that the camp and non-camp groups do not behave differently prior to the summer holiday between grade 8 and 9, at least not academically. An important worry is that the camp boys are more motivated to change their behavior than those who do not participate. Table A.10 shows the effects on the national reading score in grade 8, which is measured after the boys have applied for the summer camp, but prior to the summer camp. Thus, if motivation is driving the main effects we would expect to see significant results in this table. However, all estimates are insignificant and close to zero. To further investigate this change in behavior I construct similar placebo tests using absence, socio-emotional skills, and school well-being. Tables A.11 and A.12 show if there are any differences from grade 7 to grade 8. This is particularly interesting because behavior changes prior to summer camp could bias the ATT estimates. Absence data is measured throughout the whole school year whereas socio-emotional skills and school well-being are measured in the spring similar to the national reading test this survey is conducted at the end or after the application deadline, but prior to notification of enrollment at the summer camp. The tables show no significant evidence of increased motivation with the individual fixed effect estimates being negative and statistically insignificant. This indicates no systematic change between the two groups leading up to the camp, which supports the main results.

6. Results: Individual vs group mentoring

In this section, I estimate the effect of the change from an individual to a group mentoring strategy in the follow-up program. As

a first indication, I present a visual illustration of the difference-in-differences results from the two school years with individual mentoring against the later three years with group mentoring. Overall, Fig. 2 shows larger effects under the group mentoring follow-up program across all outcomes. The effect of program participation on the average continuous assessment mark is positive and significant in both mentoring regimes. However, the effect nearly doubles from the implementation of group mentoring. Interestingly, during the years of the individual mentoring program, there are limited effects on the personal and social readiness assessment whereas the introduction of group mentoring substantially increases these effects to 19 and 15 percentage points, respectively with the outcomes being measured halfway through the follow-up program. For academic readiness, the difference-in-differences estimates are significantly different from zero under both the individual and group mentoring follow-up program with the latter being slightly larger. These differences in readiness assessments in different dimensions are also detectable in the overall readiness assessment, in which the difference-in-difference results increase from 12 percentage points to 22 percentage points from changing the individual mentoring program to a group mentoring program during the follow-up year.

The obvious question is whether these effects are statistically significant from each other. Therefore, I investigate the consequences of the follow-up mentoring strategy using the triple differences model described in Eq. (2). Table 2 presents the triple differences estimates, i.e. the effect of substituting the individual mentoring program with a group mentoring program. The first column shows that the effect of program participation on the average continuous assessment mark is 8% of a standard deviation larger when the follow-up program consists of group mentoring. The difference is, however, not significantly different. For the educational readiness assessment, the table shows overall positive results with the effects on the personal and social readiness assessment being statistically significant. Substitution from the individual mentoring follow-up program to a group mentoring follow-up program increases the fraction of boys who are assessed personally ready by 15 percentage points and socially ready by 16 percentage points. These results imply that, in terms of personal and social readiness, camp/mentoring boys reduce the gap to the average boy by up to 60%. This is an important finding because a major problem with intensive learning camps is the transition back into the local school

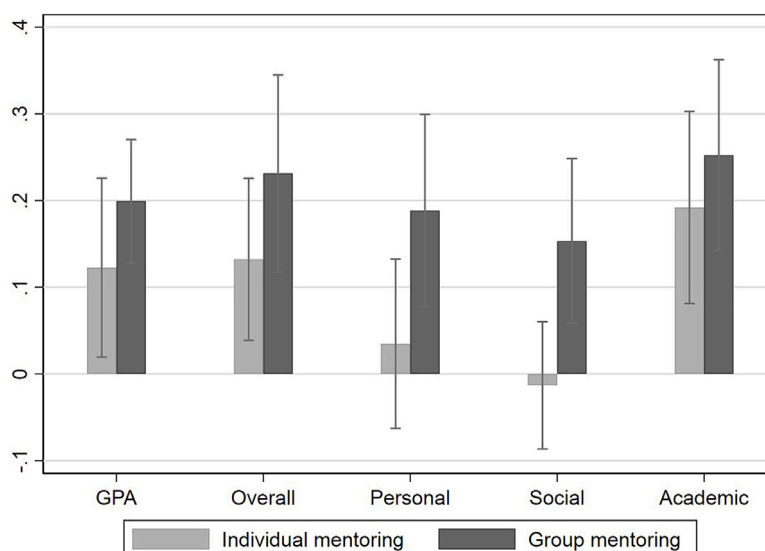


Fig. 2. Difference-in-difference estimates by follow-up program.

Notes: This figure presents the difference-in-difference effects of participating in the summer camp on outcomes separated by individual (2015–2016) and group (2017–2019) mentoring with 95% confidence intervals. GPA is the average assessment mark and is measured as standardized effect sizes. The remaining outcomes stem from the readiness assessment and can be interpreted as percentage points by multiplying with 100.

Table 2
Triple differences: Effects of group mentoring instead of individual mentoring.

	GPA (1)	Educational readiness assessment			
		Overall (2)	Personal (3)	Social (4)	Academic (5)
DiDiD	0.082 (0.065)	0.097 (0.076)	0.154 (0.076)	0.155 (0.063)	0.062 (0.081)
Observations	479,540	277,043	263,480	263,153	265,354
R-squared	0.565	0.362	0.268	0.261	0.418
Pupil background chars	Yes	Yes	Yes	Yes	Yes
Parental background chars	Yes	Yes	Yes	Yes	Yes
Reading abilities	Yes	Yes	Yes	Yes	Yes
Math abilities	Yes	Yes	Yes	Yes	Yes
Absence information	Yes	Yes	Yes	Yes	Yes
School well-being	Yes	Yes	Yes	Yes	Yes
Personality traits	Yes	Yes	Yes	Yes	Yes

Notes: This table shows triple difference (DiDiD) estimates by comparing camp participants attending the individual mentoring program with those attending group mentor centers in the year following the camp. Standard errors are clustered on the individual level. Missing values are imputed with the value zero and a binary indicator is added to the conditioning set. Bold (italic) numbers indicate significance at the 5% (10%) level.

environment. When returning to the local school and classroom, many boys perceive that “nothing has changed” and therefore, their school behavior is likely to revert to how it was before camp participation. However, the correct mentoring strategy in the follow-up program seems able to reduce this problem and avoid the complete fade-out in teacher-assessed personal and social competencies. This is crucial if we want to hope for long-run effects of educational learning programs on lifetime success (Kautz et al., 2014).

The critical assumption for the triple differences estimates to be causal is the parallel trends between the two difference-in-differences estimates. Entropy balancing ensures perfect balance across all covariates and the large set of robustness analyses conducted in the previous section show no indication of a violation of the difference-in-differences parallel trend assumption. Data restrictions – educational readiness assessment is only performed once in grades 8 and 9 and the continuous assessment mark for June in grade 8 is first registered in 2018 – making visualization of pre-trends impossible. However, as in the above section, I perform placebo tests with a fake treatment group and a fake treatment date. Table A.13 presents the triple differences results for the fake treatment group, which are all insignificant. Table

A.14 investigates academic pre-trends exploiting the national reading and math tests and a fake treatment date in the summer holiday between grades 5 and 6. All triple differences estimates are insignificant, indicating similar pre-trends between the two difference-in-differences models. Additionally, the selection process for the summer camp has not changed as illustrated in table A.15. Finally, in Figure A.5 I show the effects by cohort. Naturally, this increases the standard error, but it also shows that the triple difference results on personal and social readiness are not driven by one year, i.e. there is a level change after the change from individual mentoring to group mentoring in the program package. This indicates that the increased effects on personal and social competencies are caused by the change in the follow-up program.

7. Conclusion and discussion

I studied how the combination of summer camp and a one-year follow-up program targeted academically disadvantaged boys with a lack of school motivation affects school performance and readiness for upper secondary education in Denmark. I did so by utilizing how

boys' outcomes evolved from grade 8 to grade 9. Using a difference-in-differences strategy, I compared outcomes among those who participated in the two-week summer camp with mentoring to those who did not. Additionally, I investigated – exploiting a 2017 structural change in the one-year follow-up program – how substituting individual mentoring with group mentoring affected outcomes. Using a triple differences strategy, I compared the difference-in-differences estimates before and after the change in the follow-up program.

I found large, positive effects of camp/mentoring participation on the average assessment mark and on readiness for upper secondary education. Camp/mentoring participation increased the average assessment mark by 15% of a standard deviation and overall readiness for upper secondary education by 18 percentage points. Thus, the program reduced the gap between the average boy by up to 35% and between boys with similar characteristics by up to 80%. I found that the positive effect on the overall readiness assessment was mainly driven by increased academic competencies (22 percentage points), but the camp/mentoring also increased personal competencies (11 percentage points). Further, I saw an increase in the effects of camp/mentoring participation on all outcomes when using group mentoring in the follow-up program instead of individual mentoring. In fact, substitution from the individual mentoring follow-up program to a group mentoring follow-up program significantly increased the fraction of boys who were assessed as personally ready by 15 percentage points and socially ready by 16 percentage points. Thus, the change in the follow-up strategy led to a dramatic increase in the effectiveness of the program on non-academic competencies.

This study thus offers two key findings; (1) summer camps in combination with a follow-up program are effective in improving adolescent boys' school outcomes and (2) Using a group mentoring scheme in the follow-up program increases the effects. Interestingly, the results suggest that the effect on academic performance exists when intensive learning camps are combined with a one-year mentoring follow-up program. However, the type of mentoring program is important for the effect on non-cognitive skills with group mentoring being preferable. Intuitively, it makes sense that group mentoring, where such competencies automatically are in play, is more effective than individual mentoring in improving social and personal competencies. Additionally, group mentoring has the advantage of being cheaper because of the lower adult-to-boy ratio. These findings raise the question if a reduction in camp length and additional follow-up sessions with the mentor group, which would reduce the total cost of the intervention, is more cost-effective. A new similar one-week camp with a similar follow-up program was, in the autumn holiday of 2021, commenced by the same organization that arrange the summer camp under study. Future research exploiting both the summer and autumn camps may be able to shed light on this question.

A potential caveat of this study could be that the positive effect originates from the boys-only teaching and not from the summer camp itself. Briole (2021) shows that pupils benefit from similar peers and find that increasing the classroom with 5 girls leads to a decrease in boys' test scores by 1.6% of a standard deviation. With Danish school classes often being split 50/50 and an average class size of 21 this is a positive effect on 3.2% of a standard deviation for at least a full year of only-boys classes compared to two weeks of boys-only teaching in this study. Thus, with the small effect compared to this study and the limited time frame, it is unlikely that the effect sizes of this camp/mentoring package originate from boys-only teaching.

Kraft (2020) argues that, from a policy perspective, the effect sizes are important, but more important is the effect relative to program costs. Furthermore, he argues that learning programs are not as relevant if they cannot be taken to scale with high fidelity. This camp/mentoring program cost approximately 5,000 USD per participant, which (Kraft, 2020) defines as being at the lower end of a high-cost intervention. However, when accounting for the degree of disadvantage facing this group of adolescent boys, which the literature has shown is difficult

to affect, the effects of this camp/mentoring program appear quite impressive. There is naturally a limited number of participating boys at each camp, but the comprehensive curriculum and fixed structure should be an advantage when scaling and conducting several camps simultaneously.

Overall, the findings in this study are important for policymakers because they show that a summer camp at least in combination with a mentoring program can be a relevant and effective tool for improving academic performance for disadvantaged boys who have fallen behind academically. Furthermore, with the large impact on readiness for upper secondary education, the summer camp has the potential to become a key instrument in making pupils ready for upper secondary education. Naturally, the cost of 5,000 USD per participant has to be assessed against the benefits of camp participation. For example, the reduction in cost to society of boys' not attaining secondary education or delaying it. Finally, the many disruptions to education caused by e.g. the Covid-19 pandemic flattened the learning curve and increases the socioeconomic learning gap. Di Pietro et al. (2020) show that these disruptions to a greater extent affect disadvantaged boys by stopping their learning process and in fact decrease their cognitive abilities. Utilizing summer camps as a remedial educational program could potentially bridge the learning gap and improve academic outcomes for at-risk boys.

The scalability of the program is a limitation in this study, as it could prove to be challenging. The summer camp is annually organized by dedicated teachers, adults, and an organization committed to enhancing the life prospects of the boys, and is held at a fixed location. Further investigation is needed to determine if the program's effectiveness is maintained when expanded to a larger scale. Additionally, the program targets boys and the analysis provide no insight into its effect on girls. Thus, future research is needed on the effects of camp/mentoring participation for disadvantaged girls. Finally, a methodological limitation of this study is its reliance on a difference-in-differences strategy, which assumes parallel trends and restricts the ability to examine long-term effects, such as educational achievement, future income, employment history, etc. The next step would be to conduct a randomized trial of the summer camp, which would allow for extended follow-up and demand fewer stringent assumptions for identification.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

I gratefully acknowledge discussions and helpful comments from two anonymous referees, Marianne Simonsen, Michael Rosholm, Maria Knøth Humlum, Mette Ejrnæs, Joshua Goodman, and participants at multiple seminars and workshops in Aarhus, Bergen, and online.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.econedurev.2023.102370>.

References

- Abadie, A. (2005). Semiparametric difference-in-differences estimators. *Review of Economic Studies*, 72(1), 1–19.
- Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality psychology and economics. In *Handbook of the economics of education* (pp. 1–181).
- Andersen, F., Mølgaard, L., Wåst, T., & Nørgaard, L. (2019). *Intensiv læring - relationer, mestingsoplevelser og transfer i det nye pædagogiske landskab*. Frederikshavn: Dafolo A/S.
- Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2019). Family disadvantage and the gender gap in behavioral and educational outcomes. *American Economic Journal: Applied Economics*, 11(3), 338–381.
- Battistin, E., & Schizzerotto, A. (2019). Threat of grade retention, remedial education and student achievement: Evidence from upper secondary schools in Italy. *Empirical Economics*, 56(2), 651–678.
- Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics*, 27.
- Blundell, R., Dias, M. C., Meghir, C., & van Reenen, J. (2004). Evaluating the employment impact of a mandatory job search program. *Journal of the European Economic Association*, 2(4), 569–606.
- Briole, S. (2021). Are girls always good for boys? Short and long term effects of school peers' gender. *Economics of Education Review*, 84, Article 102150.
- Cooper, H., Charlton, K., Valentine, J. C., Muhlenbruck, L., & Borman, G. D. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development*, 1–127.
- Di Pietro, G., Biagi, F., Costa, P., Karpiński, Z., & Mazza, J. (2020). *The likely impact of COVID-19 on education: reflections based on the existing literature and recent international datasets: Tech. rep.*, LU: Publications Office of the European Union.
- Freier, R., Schumann, M., & Siedler, T. (2015). The earnings returns to graduating with honors — Evidence from law graduates. *Labour Economics*, 34, 39–50.
- Gruber, J. (1994). The incidence of mandated maternity benefits. *The American Economic Review*, 84(3), 622–641.
- Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1), 25–46.
- Hainmueller, J., & Xu, Y. (2013). ebalance: A stata package for entropy balancing. *Journal of Statistical Software*, 54(1), 1–18.
- Hattie, J., & Yates, G. C. R. (2013). *Visible learning and the science of how we learn*. London: Routledge.
- Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, 64(4), 605–654.
- Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *The Review of Economics and Statistics*, 86(1), 226–244, Publisher: MIT Press.
- Kautz, T., Heckman, J. J., Diris, R., ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success: Working paper 20749*, National Bureau of Economic Research, Series: Working paper series.
- Kim, J. S., & Quinn, D. M. (2013). The effects of summer reading on low-income children's literacy achievement from kindergarten to grade 8: A meta-analysis of classroom and home interventions. *Review of Educational Research*, 83(3), 386–431.
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241–253.
- Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research*, 76(2), 275–313.
- Mariano, L. T., & Martorell, P. (2013). The academic effects of summer instruction and retention in New York city. *Educational Evaluation and Policy Analysis*, 35(1), 96–117, Publisher: SAGE Publications Sage CA: Los Angeles, CA.
- Murnane, R. J. (2013). U.S. High School graduation rates: Patterns and explanations. *Journal of Economic Literature*, 51(2), 370–422.
- OECD (2013). *Education at a glance 2013: OECD indicators*. Paris: OECD Publishing.
- Olden, A., & Møen, J. (2020). *The triple difference estimator: NHH discussion paper*, (p. 26).
- Petrosino, A., Turpin-Petrosino, C., Hollis-Peel, M. E., & Lavenberg, J. G. (2013). 'Scared Straight' and other juvenile awareness programs for preventing juvenile delinquency. *The Cochrane Database of Systematic Reviews*, (4), CD002796.
- Rambøll (2019). *Kvalitativ undersøgelse af Plan T, DrengeAkademiet og Københavner-Akademiet: Tech. rep.*, Egmont Fonden and Bikubenfonden.
- Rosholm, M., Paul, A., Bleses, D., Højen, A., S. Dale, P., Jensen, P., M. Justice, L., Svarer, M., & Calmar Andersen, S. (2021). Are impacts of early interventions in the Scandinavian welfare state consistent with a Heckman curve? A meta-analysis. *Journal of Economic Surveys*, 35(1), 106–140.
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68–78, Place: US Publisher: American Psychological Association.