I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14136

The Importance of Peer Quality for Completion of Higher Education

Maria Knoth Humlum
Mette Thorsager

FEBRUARY 2021

I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14136

The Importance of Peer Quality for Completion of Higher Education

Maria Knoth Humlum
Aarhus University and IZA

Mette Thorsager
VIVE - The Danish Center for Social Science Research and Aarhus University

FEBRUARY 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Importance of Peer Quality for Completion of Higher Education*

Using detailed Danish administrative data covering the entire population of students entering higher education in the period 1985 to 2010, we investigate the importance of a student's peers in higher education for the decision to drop out. We use high school GPA as a predetermined measure of student ability and idiosyncratic variation in peer composition across cohorts within the same education and institution. Our findings suggest that peer ability is an important determinant of students' drop out decisions as well as later labor market outcomes. Overall, we find that a one standard deviation increase in peers' high school GPA reduces the probability of dropping out by 4.6 percentage points. This number masks considerable heterogeneity by level and field of study. Allowing for a more flexible specification, we find that low quality peers have adverse effects on the probability of dropping out while high quality peers have beneficial effects. These effects are more pronounced for lower ability students.

JEL Classification: I21, I24

Keywords: dropout, peer effects, peer quality, higher education

Corresponding author:
Maria Knoth Humlum
Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
8210 Aarhus V
Denmark
E-mail: mhumlum@econ.au.dk

* Maria Knoth Humlum gratefully acknowledges support from SparNord Fonden. We thank Nina Smith, Tor Eriksson, Fane Groes, Alexandra Spitz-Oener, and seminar participants at the PhD workshop in Stavanger 2018 for comments and suggestions.

1 Introduction

In the OECD countries, about 31 percent of the students that enroll in tertiary education later drop out (OECD, 2009). Policy makers are encouraged to reduce dropout while ensuring that student dropout does not lead to increased inequalities in educational attainment across students with differential background, for example, in terms of socioeconomic characteristics (OECD, 2008). It has also been highlighted that student composition may be an important factor in explaining increasing dropout rates in the United States (Bound et al., 2010), making it important for policy makers to understand the relevant mechanisms. Policy makers often apply policy levers that affect the student composition in higher education, for example through entry requirements and quotas. When considering the implementation of such policies, it may also be relevant to consider whether a given policy has implications for student composition and as a result possibly student dropout behavior.

The existence of peer effects has been analyzed in many different educational contexts. In higher education, the evidence is characterized by originating from various more or less selected samples of the general population of students in higher education: freshmen at Dartmouth college (Sacerdote, 2001), undergraduates in economics and business at the University of Amsterdam (Booij et al., 2017), freshmen at the United States Air Force Academy (Carrell et al., 2013), and first-year students in a public Italian university (Brunello et al., 2010) are just some of the specific groups of students where peer dynamics have been investigated.

It is plausible that peer dynamics in higher education differ depending on the context, for example, the field of study (Brunello et al., 2010), but given the scattered coverage of the existing peer effect studies, it is hard to obtain a unified pattern of the effects. We remedy this by using rich administrative data covering the universe of enrollments in higher education in Denmark in the period 1985-2010. Our sample thus consists of more than one million separate enrollments. The universal coverage and large sample size allow us to compare effects across various subgroups of interest, such as field of study and selectivity of the education program, and to allow for nonlinear peer effects. Identification of peer effects in this setting relies on idiosyncratic variation within institutions and education programs over time. To the extent that peer effects may differ across sub-populations, our approach will allow us to estimate peer effects for different subgroups in the same institutional context. Our focus on entire cohorts of college entrants ensures a larger degree of external validity than studies with a more narrow focus on a specific sub-population.

Finally, we leverage the large administrative data sets to make two additional contributions. First, we estimate the effects of peer quality on later labor market outcomes. Second, we assess the existence of potential non-linearities in peer effects in a model where peer effects depend on the distribution of peers' and own ability. While our baseline measure of ability captures ability relative to the entire high school cohort, we also employ a measure of ability that captures ability relative to actual peers (in the same education and institution). This allows for a comparison of results based on absolute versus relative ability measures.

We focus on the effects of peer ability on the students' decision to drop out of higher education which is highly relevant for policy makers. In Denmark, dropout rates in higher education are about 30% and reducing dropout rates is a continuing concern.[1]

The academic ability of peers may influence the dropout decision through two pri-

[1] An official aim of the Danish government is that 50% of 30-year-old individuals should have completed a higher education.

mary channels. First, peer ability may affect a student's own ability which in turn may decrease the probability of dropping out. For example, better peers may help improve the understanding of curriculum through interactions among students. This can happen as a result of discussions or questions during lectures, classroom seminars or even outside of teaching. Second, peer ability can affect other dimensions such as the behavior of a student. For example, students may change the amount of effort they exert with respect to their studies, or change the nature of their study habits by observing peers with high cognitive and non-cognitive skills. It is important to keep in mind that student dropout may occur for many different reasons and the decision to drop out may be optimal – not only from the perspective of the individual, but also from the perspective of society.

This paper investigates the importance of peer ability for the decision to drop out of higher education. We take advantage of rich administrative data on the universe of students enrolling in higher education in Denmark in the period 1985-2010. The identification strategy relies on idiosyncratic variation across cohorts within education-by-institution cells. We use high school GPA as a predetermined measure of student quality. First, our empirical study applies the linear-in-means model to explore the varying peer effect estimates across several educational contexts. Next, we exploit the availability of the full distribution of high school GPAs to estimate heterogeneous non-linear peer effects. We find that a one standard deviation increase in peers' high school GPA reduces the probability of dropping out by 4.6 percentage points. In subsequent analyses, we find that there is considerable heterogeneity in the effects, for example by level, field of study, and the selectivity of the education program. In a specification that allows for nonlinearities and interactions between own and peer ability, we find that low quality peers have adverse effects on the probability of dropping out while high quality peers have beneficial effects. Results show that these effects are strongest for lower ability students.

The paper is organized as follows. Section 2 provides a brief overview of the relevant literature. The institutional settings are briefly described in Section 3. In Section 4, the data is presented along with some descriptive statistics. The empirical strategy is introduced in Section 5. Section 6 discusses and evaluates the proposed identification strategy. This is followed by the presentation of the results in Section 7. Section 8 concludes.

2 Previous Literature

The literature on peer effects is growing rapidly and many studies exist on peer effects in education in general.[2] In this section, we review selected papers where the focus is on peer effects in higher education. One of the main challenges in the peer effects literature is to find exogenous variation in peer composition. In higher education, random assignment of students to housing units (Sacerdote, 2001; Zimmerman, 2003; Stinebrickner and Stinebrickner, 2006; Foster, 2006; Brunello et al., 2010; Griffith and Rask, 2014) and random allocation to classes or groups (De Paola and Scoppa, 2010; Brodaty and Gurgand, 2016; Feld and Zölitz, 2017; Booij et al., 2017) have been used as plausible identification of peer effects. The resulting estimated effects are generally small and not always statistically significant. The random allocation of peers provides a credible identification strategy; however, Angrist (2014) argues that the roommate strategy suffers from weak instrument bias.

[2]For more general surveys of empirical evidence on peer effects in education, see Epple and Romano (2011); Sacerdote (2011, 2014).

Sacerdote (2001) and Zimmerman (2003) are the first to exploit the random allocation of students to roommates to study peer effects among college roommates. They find no linear impact of roommate ability on first-year college GPA. However, with a non-linear specification Sacerdote (2001) reveals beneficial impacts from having a high-ability roommate, and Zimmerman (2003) finds that low-ability roommates are harmful. Other studies applying the roommate strategy find that peer effects are limited to female students (Stinebrickner and Stinebrickner, 2006), smaller colleges (Griffith and Rask, 2014), or students in hard sciences (Brunello et al., 2010).

Another strand of the literature exploits the random allocation of students to teaching classes or study groups to estimate peer effects among students sharing the same teaching environment. De Paola and Scoppa (2010) use data from the University of Calabria. The randomly generated peer groups in compulsory courses during First Level Degree are used to instrument teaching classes during the Second Level Degree. They find that sharing teaching classes with high-ability students significantly improves exam grades. Similar results are found by Feld and Zölitz (2017) in their empirical study using random assignment of students to within-course sections at Maastricht University. Other studies exploiting random allocation to teaching classes find that peer effects are gender-specific. Fischer (2017) finds that women are less likely to graduate with a STEM degree when grouped with higher ability peers, whereas Ficano (2012) finds that male students' academic achievement is positively affected by the ability of their male peers.

Meanwhile, no such effects are found by Brodaty and Gurgand (2016). Their study uses data from an elite French university where all teaching take place in small classes to which students are randomly allocated. Even though students share all their teaching, no significant impact from high-ability peers on student performance is detected when accounting for teacher effects. Feld and Zölitz (2017) also account for teacher effects, however, their estimate of peer effects remains statistically significant. There is no immediate explanation to reconcile the difference in findings of Brodaty and Gurgand (2016) versus Feld and Zölitz (2017) and De Paola and Scoppa (2010), except that these studies are all based on data from single universities, and peer effects may differ across institutions as shown by Griffith and Rask (2014).

When random allocation of peers is not generally feasible, other approaches have been used that take advantage of the data structure. For example, one can exploit the natural variation in average ability across years within institutions. Arcidiacono and Nicholson (2005) use this approach to estimate peer effects among the universe of medical students graduating from US medical schools. They include medical school fixed effects and thus use variation across years in within-school average ability to identify the effect of peer ability on the exam results. They find no statistically significant effects.

Arcidiacono et al. (2012) present a new empirical method operating through unobserved fixed effects. Under a set of assumptions, they estimate peer effects from a linear combination of individual fixed effects. Hence, their approach accounts for unobserved abilities. Their approach is applied to data from the University of Maryland, and they find small positive peer effects on grades.

Few studies have explored whether the ability level of peers has impacts on the decision to drop out from higher education. Luppino and Sander (2015) study the impact of peer quality on attrition from sciences. They find that students attending campuses with stronger peers in sciences are far less likely to graduate with a science degree. Fischer (2017) finds a similar result, however, the effect is limited to female students. Meanwhile, stronger peers in non-science courses increase graduation rates in both science and non-science majors. The authors argue that these pattens are due to positive peer effects on performance in non-science courses and peer competition in science courses. In a similar framework Arcidiacono et al. (2016) show that better matches of own ability to that of peers are important for persistence in STEM fields among minority students. Booij et al. (2017) find that low- and medium-ability students have lower dropout rates when the composition of tutorial groups switches from mixing to three-way tracking in a sample of undergraduate students in economics. Meanwhile, they find no evidence that high-ability students are affected by the quality of their peers.

In conclusion, the results from the peer effects literature are mixed. Many studies, especially those focusing on elite schools find only very small or no peer effects. The studies are mainly based on students from only one school or a few schools limiting the possible heterogeneities that can be explored and the external validity of the conclusions.

3 Institutional Setting

Denmark is a Scandinavian welfare state, and higher education is generally publicly provided and tuition free. In addition, the Danish state provides generous student grants. Upon completion of high school, a student can apply for a slot in higher education (university or professional bachelor's programs).[3] A high school degree is a prerequisite for university programs and for most professional bachelor's programs.

Admission to higher education is mainly based on the grade point average from high school. Applicants submit an application to the Coordinated Enrollment System where they can rank up to eight education-by-institution programs. Slots are allocated using a deferred acceptance mechanism. The GPA cutoff for each program will depend on the number of applicants, their GPA, their ranked list of programs, and the number of slots. Students are offered one slot. Students do not know the GPA cutoff in advance and therefore neither the ability of their fellow students.[4]

It is also possible to apply for higher education without a high school degree. There is a separate application process for students who want to apply for a slot based on other criteria than the high school GPA. Students are assessed based on other criteria such as non-high school degrees, for example, vocational education, and prior work experience. The institutions individually set the number of students admitted in this way.

The teaching structure at higher educations is specific to the education program. In general, most education programs use a combination of lectures for the entire student cohort, and smaller classroom-based seminars with 20-30 students. In most education programs, the cohort of students starting at the same time follow the same set of courses for the first couple of years of their education. At the universities, students typically enroll in a bachelor's program with a duration of three years and subsequently enro in a master's program with a duration of two years. The professional bachelor programs at the university colleges have a duration of 3-4 years. For the remainder of the paper, we will refer to the professional bachelor enrollments as the college sample, and university enrollments as the university sample. We have grouped all education programs into the four fields: STEM, Social Science, Arts, and Health. Table A.1 shows a subdivision of the field of study groups, and gives examples of specific education programs.

[3] Higher education in Denmark consists mainly of university and professional bachelor's programs. There are also some shorter programs which we do not include in the present description or analysis.

[4] Humlum et al. (2017) provide a more detailed description of the Coordinated Enrollment System and document the variation in GPA cutoffs over time.

4 Data

We exploit information on the universe of higher education enrollments in the period 1985-2010 in Denmark. We link several administrative registers to obtain a rich data set with background information about enrolling students, students' peers and students' parents. For each enrolling student, we have information about the institution and the education program. Usually, an education program, for example, Medicine, will be supplied by more than one institution. Students can be followed over time such that we observe whether students eventually drop out or complete the education program they enroll in.

4.1 Definition of Peer Groups

It is not trivial to define a relevant peer group for a given student. The definition of peer groups is widely discussed in the literature, and several different definitions have been applied. Within the literature of peer effects in higher education, there has been a particular focus on students sharing housing units. As pointed out by Stinebrickner and Stinebrickner (2006), flatmates have the potential to affect time-use and study habits. Hence, residential peers and their effects on student performance are of interest. However, the flatmate-definition of peer groups is not capturing the likely peer effects present among students sharing a teaching environment. Also, survey evidence from Denmark shows that 40% of dropouts list being lonely in the educational program as a reason for dropping out, UFM (2018). This suggests that students within the same educational programs constitute important peer groups.

We define students that enroll in the same education program in the same institution in the same year as a peer group. In other words, peer groups are defined based on education-by-institution cells in the year of enrollment. A smaller fraction of education programs has new intake of students twice a year (September and February). For these education-by-institution cohorts, the peer group is separated according to when enrollment occurs. Students who enroll in an education program at a given institution in the same year will generally take the same courses at the same time. Depending on the size of the education program at a given institution, these students may have tighter or looser connections. Generally, students may share all of their classes, a subset of their classes or none of their classes. Students in smaller education-by-institution cells are more likely to share all of their classes. Examples of education-by-institution cells are Economics-Aarhus University, Economics-Copenhagen University, Nursing-University College Copenhagen, Nursing-VIA University College Aarhus. We observe on average 579 education-by-institution cells per year. The median (over education-by-institutions) group size is 28 for university programs and 47 for college programs.

The size of the peer groups varies substantially across education programs. We have restricted the sample to include peer groups with a minimum size of four students i.e. minimum three peers. As a consequence of the peer group definition, a number of popular education programs have very large peer groups. In the largest cohort, there are 1,924 students in a peer group. Most likely, not all students in the largest peer groups represent a peer with important influence, but we are not able to observe smaller peer groups such as study groups or classrooms. Even if we could identify teaching classes, it is not obvious whether this information would improve the peer group precision. It is likely that students form smaller subgroups within classes or cohorts unobservable to the researcher. In addition, many education programs, provide shared lectures for the entire cohort of students. Hence, using the education-by-institution cohort of students to define the peer group captures peers that students interact with by sharing a teaching environment.

4.2 Measure of Peer Ability

Measuring peer effects is complicated by peers simultaneously affecting each other. This is commonly referred to as the problem of reflection bias (Manski, 1993). Following Lavy et al. (2012) who study peer effects in English secondary schools, we use a predetermined measure of ability - high school grades that are determined prior to enrollment in higher education. At high school (HS) graduation, grades from courses completed in high school are averaged to a high school GPA. The availability of HS GPA varies for graduates across the types of high school program.[5]

High school GPAs are standardized within year of high school graduation to account for potential grade inflation. We calculate the peer group average of the standardized values of GPA omitting own HS GPA to obtain the leave-one-out average. Also, for each year of high school graduation, we use the distribution of high school GPAs to calculate the decile of an individual's GPA and define indicators of which decile a student's grade is in. We then use these indicators to calculate the share of peers in each decile (again adjusting for own value).

Not all students in higher education have completed a high school degree, and high school degrees completed before 1978 are not observed perfectly. Therefore, 29% of students in our sample do not have an (observable) high school degree recorded. For students without an observable value of own HS GPA, we impute the peer ability measure with the peer group average HS GPA (of observable values) instead of the leave-one-out-average. In the regressions, own HS GPA is set to zero if missing, and a dummy variable is included to control for the missing HS GPA score.

4.3 Estimation Sample

In the period 1985-2010, we observe 1,203,232 enrollments in higher education distributed across 822,664 individuals. In Denmark, it is common to enroll in more than one higher education. For example, a student may decide to drop out and enroll in another higher education. Since we do not restrict the sample to first-time enrollments, the number of enrollments is higher than the number of individuals.[6] We delete 9,654 observations where the peer group consists of three students or less. We exclude 2,847 observations where there is no observable value of peer group ability. Next, we exclude 4,025 enrollments where there is only one observable high school GPA for the peer group. We observe 2,616 enrollments in an education-by-institution combination that only exist for one particular cohort. Since we rely on year-to-year variation within education-by-institution cells, these observations are dropped from the sample. The estimation sample comprises 1,184,090 enrollments (813,337 individuals) across 1,181 education-by-institution combinations.

4.4 Summary Statistics

Table 1 shows key descriptive statistics for the estimation sample.[7] The first column shows the average values of characteristics for the sample of university students while the second column shows the averages for college students. 51% of the university students are women, while the female share of college students is 67%. Women are overrepresented

[5] HS GPA is available for STX and HF graduates from 1978 and forward. HTX and HHX graduates have observable HS GPA from 1999 and forward.

[6] Restricting the sample to first-time enrollments reduces sample size substantially, but yields similar results, cf. Appendix Table A.2. Statistical significance is generally reduced for the sample of first-time enrollments.

[7] See summary statistics separately for field of study in Table A.3 (university) and Table A.4 (college).

in higher education in Denmark compared to men. This is driven by high proportions of female students in Arts and Health. Students are about 25 years old on average when they enter higher education. The median age at entry is 23 years.[8] The students generally tend to have favorable socioeconomic background characteristics in terms of parental education. 45% of university students have parents with a high education level. The average dropout rate from university is 33% and varies from 21% in Health at university to 42% in Arts at university. College students have lower rates of dropout on average. The median (over students) number of peers in university is 123 reflecting that the peer group is defined in relatively broad terms as the entire entry cohort of the education in a particular institution. Peer groups are slightly smaller in college programs.

[INSERT TABLE 1]

Table 2 describes the peer ability distribution across university and college.[9] There is substantial variation across the two education programs. In total 5% of university students are in the bottom 20% (as ranked by their high school GPA in their high school graduation cohort) and 27% are in the top 20%. College programs enroll more students from the bottom 20%, and fewer from the top 20%. In general, the distribution of students' ability is skewed to the left for university programs, and to the right for college programs. University Health programs vary substantially from other university fields. 46% of university Health students are in top 20%, while the corresponding numbers are 27% in STEM, 25% in Social Science and 23% in Arts. There is also substantial variation across college fields; in college Social Science programs, 13% of the students were in the top 20% and 11% were in the bottom 20% whereas in college Health 17% of students were in the bottom 20% and only 4% of students were in the top 20%. It is clear from Table 2, that university students are characterized by higher levels of academic ability. Also, GPA is missing for 23% of university enrollments, compared to 38% of college enrollments. These differences suggest that peer effects may vary across level of education.

[INSERT TABLE 2]

4.5 Dropout and Ability Trends

Dropout rates vary across time and across education programs. Figure 1 plots the differential dropout trends for first-time enrollments in university and college programs. Rates of dropout from university programs have been decreasing over the sample period. Meanwhile, dropout rates from college programs have been increasing since 1995. Yet, during the period 1985-2010, the average dropout rate from college has been lower in each of the years compared to university dropout rates.

Figure 1 also plots the average academic ability (measured by HS GPA) of first-time enrollments across the sample period. The average ability of university students varied substantially in late 1980s and early 1990s. In 1988, a large reform of the Danish Government Grant Policy took place where the student grant amount more than doubled. The reform had small impacts on enrollment rates (Nielsen et al., 2010). If the supply of study slots was fixed in short run, this may explain the small effect on enrollment rates, and the spike in academic ability of university students in 1989 reflects the higher

[8] Danish students are relatively old when they enroll in higher education, OECD (2018) and Humlum (2007). According to OECD (2018), the median age at first entry into tertiary education in Denmark is 21 whereas the 80th percentile is 26. For first-time enrollments, the median age at entry is 22 years in our sample.

[9] See footnote 7.

probability of enrollment for students with higher HS GPA scores. In more recent years, the ability of university enrollments has stabilized, but still displays variation from year to year. Among college enrollments, there has been an overall downward trend in the average ability of college students.

[INSERT FIGURE 1]

5 Empirical Strategy

Peer composition in higher education is not random given that students sort into education programs and institutions based on numerous unobservable characteristics. To avoid the selection bias created by the sorting of students, our empirical approach exploits the natural variation in cohort composition within institutions and education programs across years to identify peer effects. This approach allows for students to sort into education-by-institutions based on average education-by-institution attributes, but not based on cohort-specific deviations. This type of strategy has been applied to identify peer effects in several related contexts; Danish adolescents, Bertoni et al. (n.d.) and Brenøe and Zölitz (2020), Norwegian adolescents, Black et al. (2013), U.S. medical school students, Arcidiacono and Nicholson (2005), and Texan school children, Hoxby (2000). The key identifying assumption is that the variation in peer composition is exogenous after controlling for education-by-institution fixed effects. The greatest threat to this identification assumption is the potential presence of education-by-institution specific trends affecting both the peer composition and the dropout decision. To capture such correlations, we include education-by-institution linear time trends.

Consider an individual i enrolled at education-by-institution s in year t. Let y_{ist} be an indicator of dropout. The structure of the model is as follows:

$$y_{ist} = X_{ist}\beta + g(A_{ist}) + \bar{Z}_{ist}^{-i}\delta + f(A_{ist}^{-i}) + \mu_t + \theta_{st} + \gamma_{st} \cdot t + \epsilon_{ist} \tag{1}$$

where X_{ist} is a vector of individual background characteristics, $\bar{Z}_{ist}$ is a vector of peers' background characteristics, and $g(A_{ist})$ and $f(A_{ist}^{-i})$ are functional forms of own and peers' high school achievement, respectively. μ_t captures cohort fixed effects while θ_{st} and γ_{st} control for education-by-institution fixed effects and linear time trends, respectively. The individual characteristics are observed at age 16 and include controls for gender, age at enrollment, indicator of non-Western descent, number of siblings, highest parental education, and parental income. The peer characteristics are average values of the individual characteristics within a peer group (omitting own value of the variable in calculation of the mean). We also include controls for the number of prior enrollments, cohort size, high school type completed, and year of high school graduation.

We remove any potential effects stemming from old peers, that is peers from a student's high school that also enter the student's peer group in higher education. That is, we calculate the ability measure for peers from the same high school graduation cohort, and add this as a control in our estimation.

For the standard linear-in-means model, $f(A_{ist}^{-i}) = \lambda \bar{A}_{ist}^{-i}$, λ is the main parameter of interest. Assuming the key identifying assumption is not violated, λ captures the causal effect from the academic ability of peers on students' dropout decision. The reduced-form estimation of λ possibly captures both contextual and endogenous peer effects following the classification of Manski (1993). In other words, the estimated effect of peer quality on a student's dropout decision may reflect how student dropout behavior is affected by both peer characteristics (a contextual effect) and peer behavior (an endogenous effect).

Very few studies have attempted to separate these. Such an attempt requires additional assumptions and exogenous variation in both contextual and endogenous characteristics of peers, see, for example, Bramoullé et al. (2009). Our analysis will document the existence of peer effects without making further attempt to separate the effects.

The idea behind this strategy is to exploit the lack of Law of Large Numbers which is likely prevalent in small cohorts. Even in larger cohorts we may observe cohort-to-cohort variation in average ability of peers for a number of reasons. First, the individual education institutions have the discretion to decide the relative admission numbers of quota 1 and quota 2 applicants. Quota 2 applicants are assessed on other characteristics (such as previous work experience, stays abroad etc) if they do not qualify based on high school GPA. Hence, quota 2 applicants are generally of lower academic ability meaning an increase in quota 2 intake lowers the average academic ability of a cohort. Second, the demand for higher education or specific education programs varies with the size of youth cohorts, fluctuating returns to education, or trends in education preferences. Due to the deferred acceptance mechanism, increasing demand likely increases the average ability of a cohort. Finally, the supply of available slots can be regulated, and new study programs may open.

6 Evaluation of the Identification Strategy

Our identification of credible peer effects hinges on two things. First, we need sufficient variation in the peer ability measure to identify peer effects. In Table 3, we have decomposed the education-by-institution mean ability to describe the identifying variation in the data. Following Ammermueller and Pischke (2009), we adjust the between component for the decomposition to add to the total variance in an unbalanced panel. The upper panel of the table shows the decomposition of the variation using the raw data. The decomposition into overall, between and within variance has been performed for the full sample, and separately for the college and university samples. As one would expect, the within variation constitutes a smaller share of the total variation than the between variation. About 23% of the total variation constitutes variation within education-by-institution cells in the full sample. The within variance constitutes an even higher share in the sub-samples of university and college students. The lower panels of Table 3 demonstrate that conditioning on year fixed effects, education-by-institution fixed effects and trends successfully eliminates the between variance while a non-negligible amount of within variance remains.

[INSERT TABLE 3]

Second, changes in the peer ability measure within education-by-institutions must be uncorrelated with unobserved factors that affect student's achievement. We assess the validity of this assumption from a balancing test approach following Lavy and Schlosser (2011). We check whether the within education-by-institution variation in the peer ability measure is correlated with background characteristics of students or characteristics of the peer cohort. If the within education-by-institution variation in the peer ability measure is uncorrelated with selection into cohorts, we would expect to estimate zero correlations. Table 4 reports the results of the balancing tests from regressions of various student characteristics on the peer ability measure conditioning on education-by-institution fixed effects and trends, year fixed effects, and controls of peer characteristics. The balancing tests have been performed separately for the university and college samples.

In the sample of university students, the peer ability measure is not related to most of the observable student characteristics. The only exceptions are the birth month and cohort size. The positive association between the peer ability measure and birth month is significant only at a 10 percent confidence level, and the practical size of the association is very small. The point estimate suggests that a standard deviation increase in the ability of peers is associated with being born 0.039 months later in a year. The positive association between peer ability and the cohort size is more concerning. However, we include cohort size as a control in our estimation model.

In the college sample, we find several indications that the variation in the ability of peers correlates with background characteristics of the students. However, in all instances the practical association with the variables is very small. For example, a standard deviation increase in peer quality is associated with being 0.11 years younger, 1.06 (0.91) months longer education of the mother (father), and being born 0.055 months later in a year. These correlations may, however, be due to systematic correlations with unobservables.[10] We cannot rule out this possibility even though we include all of the balancing variables as controls in our estimation model. Overall, we have confidence that the specifications are likely to be valid, but we are cautious in the interpretation of results based on the college sample.

[INSERT TABLE 4]

7 Results

The estimation of peer effects takes the linear-in-means model as the point of departure. To relax some of the assumptions inherent in the linear-in-means model, we proceed with the estimation of more flexible specifications. Given the large sample size, it is possible to split the data along many different dimensions including education level and field, selectivity, and individual characteristics such as ability and gender.

7.1 The Linear-in-Means Model

We begin the examination of peer effects with the simple linear-in-means model. Table 5 displays the peer effect estimates and the coefficient on own ability from specifications with increasing number of control variables. Standard errors are clustered at the education-by-institution level. The peer effect estimate measures the marginal effect from increasing peers' average (standardized) high school GPA on the probability of dropout.

In column (1), we find no significant effect from peers' ability on own dropout decision in a specification with only year fixed effects and personal background characteristics included. In column (2), we add peer group averages of the personal characteristics and a control for the academic ability of "old" peers. We detect a small, but significant, peer effect on dropout once we control for selection based on observable background characteristics. However, there is a mechanical negative correlation between own HS GPA and the peer ability measure due to the leave-one-out strategy. To eliminate this source of bias, we include own HS GPA in column (3). Now, the peer effect estimate is insignificant whereas we see a strong correlation between own ability and the probability of dropout. Once we control for education-by-institution fixed effects in column (4), having academically stronger peers significantly reduces the likelihood of dropout. Finally, we add education-by-institution specific linear time trends to the model. The estimated peer

[10] Altonji et al. (2005) suggest that the degree of selection on observables is a good indicator of the degree of selection on unobservables.

effect on dropout is slightly reduced, but remains negative and statistically significant.[11]

We find that dropout decreases by 4.6 percentage points when peers' GPA increases by a standard deviation. The size of the effect is very close to the one estimated by Booij et al. (2017). They find that a standard deviation increase in peers' GPA reduces dropout by 4.2 percentage points. Meanwhile, Luppino and Sander (2015) estimate an increase in graduation rates of 8.7 percentage points among non-STEM students from a one standard deviation increase in the ability of non-STEM peers. Hence, our linear-in-means estimate of peer effects has the same order of magnitude as estimates found in previous studies.

[INSERT TABLE 5]

7.2 Exploring Heterogeneous Peer Effects

The literature on peer effects in higher education has reported mixed findings with respect to significance and sign of the effects. Variations in ability and identification strategies offer some explanation to the differences in findings. Another explanation is that institutional differences in teaching structure, size of cohorts, student composition, and course-types affect the importance of peer effects. We can exploit the comprehensiveness of our data to test differences in the importance of peer effects across various institutional characteristics.

7.2.1 Education Level and Field of Study

The first dimensions we test are university versus college and differences across field of study. University and college programs differ in the teaching structure which may give rise to differential importance of peers' ability. In general, college programs differ from university programs by being practically oriented and with time spent in internships. Carrell et al. (2009) have estimated strong peer effects in an educational setting where students spend large amounts of time with their randomly assigned peers. Hence, less time spent in classrooms and more time in internships may result in less influence from peers among college students. Also, the practical orientation of the teaching may lower the importance of peers' academic abilities on own performance in college programs.

[INSERT TABLE 6]

Column (1) of Table 6 and Table 7 report the estimated peer effects for university and college samples, respectively. Columns (2)-(5) of the tables report the peer effect estimates based on field of study sub-samples within the two education types. We find that peer effects are on average smaller in college programs. The size of peers' effect on dropout is about half of the effect estimated for university students. Yet, even greater differences are detected when comparing estimates across field of study.

Empirical studies have identified future earnings, gender-specific preferences (Gemici and Wiswall, 2014), ability sorting (Arcidiacono, 2004), parental influence (Humlum et al., 2019), and subjective expectations (Stinebrickner and Stinebrickner, 2014) as strong predictors of field choice.[12] Due to the selective field choices, there may exist systematic

[11] To investigate the possibility that peer quality potentially affects students differently depending on the stage of their studies, we have also considered more detailed measures of dropout that capture dropout within one, two and three years, respectively. Using these alternative measures yields very similar results and the estimated peer effects are similar in magnitude and statistical significance across the three dropout outcomes.

[12] Altonji et al. (2016) give a thorough summary of empirical findings on determinants of field choice in higher education.

differences in the traits of students across fields that can give rise to variation in peer influence.

Brunello et al. (2010) have proposed a theory implying that peer effects should be strongest in fields with higher marginal returns. Indeed, their empirical study finds strongest peer effects in hard sciences, and small insignificant effects in social science and humanities. Alternatively, Luppino and Sander (2015) and Fischer (2017) argue that peer effects in STEM may be negative due to a tendency of strong competitive teaching environments and rigid grading curves. Fischer (2017) finds a significantly negative impact on female STEM graduation rates from higher ability of peers, whereas Luppino and Sander (2015) show general decreases in the likelihood of STEM graduation from attending a campus with stronger peers in STEM. Since the peer effects estimated by Brunello et al. (2010) are based on roommate ability, it may not fully capture the competitive dynamics potentially present in a classroom. This may partly explain the difference of results.

[INSERT TABLE 7]

The samples applied in the empirical studies on peer effects in higher education cover a range of education fields. Most European studies apply data on students in the fields of business administration or economics (Booij et al., 2017; Feld and Zölitz, 2017; De Paola and Scoppa, 2010; De Giorgi et al., 2010). Other papers have used data on medical students (Arcidiacono and Nicholson, 2005), STEM course participants (Fischer, 2017), and mixed college samples (e.g. Brunello et al., 2010; Luppino and Sander, 2015; Arcidiacono et al., 2012; Zimmerman, 2003). In light of the theoretical prediction of variation in the importance of peers across educational fields, we would expect to find mixed results in the peer effect literature.

We exploit that students are enrolled into a specific education program, thus the decision of field of study is made at entry to university/college. We have grouped all education programs into the four fields: STEM, Social Science, Arts, and Health.

In all sub-samples the estimated effect from peers' ability on own dropout decision is negative. However, the size and significance of the effect differ across fields. In the university sub-sample, we find the strongest peer effects in STEM programs. This finding is in line with Brunello et al. (2010). Unlike Luppino and Sander (2015), we do not find peer effects corresponding to a theory of a competitive teaching environment in STEM programs. Fischer (2017) only found negative peer impacts among female students in STEM. However, later we show that the differences of results cannot be explained by gender differences nor selectiveness of study programs. One explanation may be that the cost of switching major at US colleges is much lower than in the Danish education system where students have to dropout and apply for enrollment at a new program (Malamud, 2011).

The estimated peer effects in university Social Science and Arts are smaller in size, but not statistically different from the effect in STEM programs. Similar to the European studies showing significant peer effects from samples of students in business administration or economics, we find strong peer effects on dropout in Social Science. Though, Brodaty and Gurgand (2016) find no evidence of peer effects in a sample of undergraduate economics students. However, as we will show later, the reason may be that they use a sample of elite students.

Peers' academic ability have a much smaller impact on own dropout decision among university Health students. Arcidiacono and Nicholson (2005) found no evidence of peer effects at US medical schools. They argue that the low importance of peers could be due to medical schools being highly selective, and that medical students are often a relatively mature group. Our descriptive statistics in Table A.3 show that university Health students are not on average older, but they tend to be of higher ability. Alternatively, these insignificant findings may suggest that the ability measured by high school GPA is less important in Health programs where patient care and medical skills are pivotal.

The variation in peer effects across field of study is also found among college students. Yet, the pattern differs a bit. Columns (2)-(5) of Table 7 show that college students in Social Science experience the numerically largest peer effects, although the effect is not statistically significant. While peer effects among college STEM students are statistically significant, they are small relative to those found among the corresponding university students. Again, we do not find strong evidence of peer effects in Health programs.

So far, we have documented the existence of peer effects in higher education. The ability of peers has significant impacts on the decision to drop out. Meanwhile, the size and significance of ability[13] peer effects vary across education level and field of study. Thus, the mixed findings in the empirical literature on peer effects in higher education can partially be explained by variation in the data samples with respect to education level and field of study. Generally, we find small and insignificant peer effects in Health programs and education programs with strong focus on practical skills.

7.2.2 Selectiveness of Education Program

Selectiveness of the education institutions included in a sample is another characteristic that can give rise to variation in the estimated peer effects. As argued by Stinebrickner and Stinebrickner (2006), students entering elite schools "are likely to arrive at school with strong academic ability, good study habits, and strong beliefs about the importance of college" and that these features may mitigate any influences from peers.

Higher education in Denmark is provided mainly by public institutions, and therefore our data do not allow a straightforward definition of elite schools. However, the combination of the Coordinated Enrollment System and excess demand produces very high GPA cutoffs of some education programs. We define an education program as selective if the program in more than 50% of the years being offered have enrolled cohorts with an average GPA in the top 20th percentile of all programs (within university and college).

[INSERT TABLE 8]

We have formed sub-samples based on the indicator of selectiveness and estimated peer effects separately for selective and non-selective education-by-institution units. Results are shown in Table 8 (9) for the university (college) sample. The upper panel reports peer effects from non-selective programs, the middle panel shows estimates from selective programs, and the lower panel reports a test of the difference between the two peer effect estimates. In the sample of university students, we estimate smaller peer effects in selective programs. We find particularly large differences between selective and non-selective programs in STEM. The difference in estimated peer effects between selective and non-selective Social Science programs is also relatively large, however, the difference is not statistically significant. For Arts and Health programs, we do not detect any sizable differences in peer effects between selective and non-selective programs. In the college sample, we do not observe noticeable differences between selective and non-selective education programs in the full college population. Yet, in the sub-sample of selective Social

[13] High school course-specific grades are available from 1998. These allow us to explore the relative importance of specific skills across fields. Table A.5 shows that peers' math skills is the driver of peer effects in university STEM although we lack precision to draw firm conclusions, whereas verbal skills (measured by Danish grades) is a strong driver of peer effects in university Arts.

Science college programs, the estimated effect of peer quality is positive, and the difference between selective and non-selective programs is marginally statistically significant.

[INSERT TABLE 9]

Sacerdote (2001) and Zimmerman (2003) have found only small peer effects using samples from the selective schools Dartmouth and Williams College. Brodaty and Gurgand (2016) do not find any evidence of peer effects using data from an elite French university. Meanwhile, Stinebrickner and Stinebrickner (2006) and De Paola and Scoppa (2010) find strong peer effects in samples from non-selective institutions. Yet, significant peer effects have also been identified in selective samples (Booij et al., 2017). Based on our findings, we are not able to conclude whether smaller peer effects in selective institutions support a pattern in the literature, or if variation from cohort to cohort is too small to identify the effects. However, in Section 7.3.1, we show that peer effects in selective education programs are consistently smaller when using a relative measure of peer ability.

7.2.3 Gender Differences

It is widely debated in the peer effects literature whether male or female students are most strongly affected by peers. Stinebrickner and Stinebrickner (2006), Arcidiacono and Nicholson (2005), and Fischer (2017) identify strongest peer effects for female students, whereas Zimmerman (2003), Ficano (2012) and Griffith and Rask (2014) show that male students are more susceptible to peers' ability level. Finally, Feld and Zölitz (2017) and Brunello et al. (2010) find no significant gender differences.

In panel A of Table 10 (11), we test for gender differences in peer effects for university (college) students. For male university students the dropout rate is reduced by 5.9 percentage points from a standard deviation increase in peers' ability, and the interaction suggests a 1.0 percentage point smaller (in absolute terms) peer effect for female students. The difference is not statistically significant, though. Looking across the sub-samples by field, we neither detect any significant gender differences. Among college students, we find that female students are less affected by peers compared to male students. While the gender difference is relatively large in light of the comparably small peer effect estimate, we cannot detect it in the sub-samples by field and the sign of the gender difference estimate varies across sub-samples. Hence, we do not find a unified pattern of gender differences in the susceptibility to peers.

A few papers have found that peer effects form along gender lines (e.g., Arcidiacono and Nicholson (2005), Ficano (2012), Brenøe and Zölitz (2020)). To test this, we have split the peer ability measure by gender, and included an ability measure for both male and female peers in our regression. We have performed the regression by gender, and the results are shown in panels B and C of Table 10 (11) for the university (college) sample. In the sub-sample of female university students, we find that they are as good as equally affected by both male and female peers. The square brackets report the p-value from a Wald test of the difference between the effects from male and female peers. None of the tests report a statistically significant difference. Similar results are found among male university students. In general, we do not find strong evidence that university students are differentially affected by male and female peers. If any, the sign and size of peer effects estimated from students in Health programs suggest small differences in the susceptibility to peers. Yet again, the difference is not statistically different.

[INSERT TABLE 10]

In the overall college sample, results indicate that female students are significantly less affected by their peers' ability relative to male college students. Results from the gender-specific sub-samples show in addition, that female college students are mainly affected by their female peers. Looking across sub-samples by field, we find variations in the size and sign of gender difference. Yet, there are no statistically significant differences in any of these sub-samples.

In line with Feld and Zölitz (2017) and Brunello et al. (2010), we do not find that male and female university students are differentially susceptible to peer influences. Among college students, we find some indications of peer effects forming along gender lines.

[INSERT TABLE 11]

7.2.4 Later Labor Market Outcomes

Having established effects of peer quality on the probability of dropping out, we now estimate the effects of peer quality on two labor market outcomes: employment probability and labor market earnings (1,000 DKK in 2015 prices). Both labor market outcomes are measured as the average outcome 7-12 years after enrollment in higher education. Other studies have documented effects of educational inputs on labor market outcomes, for example, Chetty et al. (2011) and Chetty et al. (2014). Table 12 shows that the effects of peer quality in higher education on later labor market outcomes are positive and statistically significant. The average employment probability is 0.57 (0.60) and average labor market earnings are 341.97 (297.70) for university (college). Therefore, a one standard deviation increase in peer quality leads to a 1.1 (2.7) percent increase in the employment probability (labor market earnings) for university students. These results document that peer quality in higher education matters not only for outcomes during enrollment such as dropout, but also for more long-term outcomes in the labor market.

[INSERT TABLE 12]

7.3 Non-linear Peer Effects

In Section 7.1 we showed the existence of peer effects in higher education while Section 7.2 showed that characteristics of education programs determine the importance. However, the linear-in-means model, applied so far, restricts the structure of estimated peer effects. First, the peer effects are assumed homogeneous for all students irrespective of their own ability level. Second, as argued by Hoxby and Weingarth (2005), from a public policy perspective the linear-in-means model is not interesting since the net effect of student allocations is constrained to zero. To loosen this assumptions, and to test for non-linearities of peer effects, we have considered a more flexible model where peer effects depend on the distribution of peers' and own ability. To do so, we use the distribution of peers' HS GPA and interact the peer ability measures with own ability level.

To begin with, we define students with a HS GPA in the bottom 20th percentile as low-achievers, and students with a HS GPA in the top 20th percentile as high-achievers.[14] We substitute, in our estimations model, the average GPA of peers with the share of low- and high-ability peers. That means, the share of peers with middle-ability (HS GPA in the 20th to 80th percentile) are left as the base category. Equation 2 summarizes our non-linear estimation model.

[14]Students' ability percentile is determined using own high school GPA and the distribution of GPA scores in the year of high school graduation.

$$y_{ist} = X_{ist}\beta + g\big(A_{ist}\big) + \gamma_1 sq^{80}_{ist} + \gamma_2 sq^{20}_{ist} + \bar{Z}^{-i}_{ist}\delta + \mu_t + \theta_s + \gamma_{st} \cdot t + \epsilon_{ist} \quad (2)$$

where $sq^{80}_{ist} = \frac{1}{N_{st}-1}\sum_{j\neq i}^{N_{st}} I_{jst[A_j>\tau_{80}]}$, $sq^{20}_{ist} = \frac{1}{N_{st}-1}\sum_{j\neq i}^{N_{st}} I_{jst[A_j<\tau_{20}]}$, N_{st} is the peer group size, and τ_{80} and τ_{20} are the 80th and 20th percentile of HS GPA, respectively.

In the following, we only show results using the sample of university students.[15] Appendix Table A.6 and Figure A.2. Panel A of Table 13 shows the results from the estimation of Equation 2. We display the parameters of interest, γ_1 and γ_2. The coefficient γ_1 (γ_2) can be interpreted as the marginal effect on dropout from increasing the share of high-ability (low-ability) peers while holding the share of low-ability (high-ability) peers constant. Hence, by reducing the share of middle-ability peers. Column (1) displays the results for the full university sample, and columns (2)-(5) show the results by field. Corresponding to the results from the linear-in-means model, we find that increasing the share of low-ability peers is harmful, whereas high-ability peers are beneficial for graduation. In line with the results found in Section 7.2.1, we find the strongest peer effects in STEM.

Next, we interact the share of low- and high-ability peers with an indicator of own ability (low, middle or high) to measure heterogeneous effects. We display the results in panel B of Table 13.

[INSERT TABLE 13]

In the full university sample, estimates suggest that persistence of high-ability students is invariant to the distributional composition of their peers' ability. Neither increases in the share of high-ability peers nor low-ability peers affect their dropout decision significantly. Meanwhile, we find that middle- and low-ability students respond significantly to the ability composition of their peer groups. In general, increasing the share of high-(low-)ability peers decreases (increases) the probability of dropout. Point estimates suggest that low-ability students reduce their risk of dropout by 1.19 percentage points when the share of high-ability peers increases by 10 percentage points. The corresponding decrease for middle-ability students is 0.99 percentage points. Meanwhile, the harmful impact from low-ability peers is much stronger. The risk of dropout for low-ability students increases by 2.88 percentage points if the share of low-ability peers increases by 10 percentage points. The pattern of the non-linear peer effects is somewhat consistent across field of study except that high-ability Social Science students appear to be influenced by the share of low- and high-ability peers. And, for these students, low-ability peers actually reduce the risk of dropout. For all other groups of students, we find negative—or statistically insignificant—effects of the share of low-ability students.

To allow for even more flexibility, we have estimated an interaction model similar to Hoxby and Weingarth (2005). In this model, nine variables measuring the share of peers in each ability decile (40th to 50th percentile is left as the base) are fully interacted with indicators of each students own ability decile. Increasing the flexibility of the model does not change the overall findings of heterogeneous non-linear peer effects. Large shares of peers in the lowest three deciles increases the risk of dropout of equally low-ability students. Meanwhile, peers in the top three ability deciles have positive impacts on persistence rates for all students. We summarize the results in Appendix Figure A.1.

In sum, the results from Table 13 and Figure A.1 suggest that low-ability students are more likely to complete a university degree if they are grouped with academically strong peers. A similar result was found by Carrell et al. (2009). However, in a later study

[15] 38% of college students have a missing HS GPA. Hence, calculating the share of low- and high-ability peers will entail large measurement errors. Nevertheless, we have estimated non-linear models using the college sample, cf.

(Carrell et al., 2013) where this finding was used to form "optimal" peer groups, peer effects on academic achievement where negative. Hence, one should be careful to base an optimal policy design from these findings since social interaction patterns may depend on the ability composition of a peer group. Middle-ability peers may play an important role in bridging the interaction between low- and high-ability students.

7.3.1 Relative Measure of Peer Ability

The heterogeneous non-linear peer effects presented in Table 13 are difficult to compare with the literature. In our definition of low- and high-ability students, we exploited the availability of the full population of high school graduates. Therefore, in our sample low-ability students are true low achievers relative to all high school graduates. In the literature, the definition of low- and high-ability varies greatly, and in addition, there is a tendency of ranking students relatively within the sample. As an example, Booij et al. (2017) define the bottom third of their sample as low-achievers. However, their sample of undergraduate economics students is generally from the top 20% of the ability distribution. Hence, our group of low achievers is presumably very different from that of Booij et al. (2017).

The use of a within-sample relative rank of students may likely be one explanation for the mixed findings of the literature with respect to non-linear peer effects. For instance, Carrell et al. (2009), Griffith and Rask (2014) and Sacerdote (2001) find that low achievers benefit from having peers of higher ability. The reverse finding has been made by Feld and Zölitz (2017), Carrell et al. (2013), and Booij et al. (2017). As these studies vary in the selectivity of their samples, the average ability level of "low-ability" students varies across the studies as well.

We have tried to use "within-sample" relative ranking of students to define low- and high-achievers. To do so, we have calculated the bottom and top 20th percentile of HS GPA of enrollments within a specific education-by-institution unit. The education-by-institution specific percentile cutoffs then determine whether a students is a low-, middle, or high-achiever. These indicators are then used to calculate the share of (relatively) high- and low-ability peers. Hence, using this definition of "low-ability" means that low-achievers are not necessarily of low academic quality, they just have a lower HS GPA relative to the other students enrolling in the same education-by-institution. Table 14 replicates panel B of Table 13 using the relative measure of peer ability.

[INSERT TABLE 14]

Interestingly, we find small variations in the estimated peer effects using the relative measure of ability. Results show a small, but statistically significant, peer effect on dropout rates of (relative) high-ability students from having more peers of relative high ability. These effects appear to be mainly driven by students in STEM, though. Among Social Science students, the susceptibility of high-ability students to the ability composition of their peers has diminished and turned statistically insignificant. Among low-ability students, we still find a harmful impact from being grouped with a large share of similar low-ability peers. Low-ability students in Health are the exception, since the significant impact shown in Table 13 has disappeared.

The differences in the estimated peer effects from using an absolute and relative measure of peer ability may arise if the ability to learn from peers depends on the relative distance of ability to peers. The relative ability distance may determine how friendships are formed. This would explain why we find significant peer effects among high ability students using the relative measure of peer ability. Alternatively, the differences between

Table 13 and Table 14 are simply due to middle-ability students - who on average are sensitive to peers - are being distributed across the relative ability groups. Nevertheless, these differences highlight that the definition of low- and high-ability of students affects the estimated importance of non-linear peer effects.

Recall, in Table 8 we found significantly smaller peer effects in selective education programs. However, the applied definition of selective education programs is potentially limiting the identifying variation in the absolute peer ability measure. We have checked the robustness of the differences between selective and non-selective education programs using the relative measure of ability. The results are shown in Appendix Table A.7. The same definition of selectivity (using the absolute measures) has been applied. The square brackets report the p-values from a Wald test of the difference between the peer effect estimates in selective and non-selective education programs. Corresponding to the results in Table 13, we find consistently smaller peer effects for relatively low- and middle achievers in selective education programs. However, peer effects among students of relatively high ability are the exception. Table A.7 suggests that increasing the share of peers of relatively high ability, in selective education programs, increases the likelihood of dropout for equally high-ability students.[16] The overall pattern of the results in Tables 8 and A.7 point to generally smaller peer effects in selective education programs.

7.4 Robustness Checks

Our findings suggest that dropout rates of students in higher education are affected by the peer ability composition, and that the importance of peer effects depends on own ability. An important feature of our analysis is the inclusion of education-by-institution-specific linear time trends. Appendix Table A.8 replicates the peer effect estimates from Tables 6 and 7 excluding education-by-institution linear time trends. The results show that the size of the estimated peer effects are smaller when controlling for linear time trends highlighting the importance of including these group-specific trends.

For the sub-sample of university enrollments our balancing test showed a small, but statistically significant correlation between our peer ability measure and the size of cohorts. Considering our peer group definition entails estimation of peer effects from very large peer groups (>500 students), one may fear that measurement error in the peer ability measure causes an overestimation of peer effects (Feld and Zölitz, 2017). We have examined whether our results are driven by large peer cohorts. Table 15 shows results from the linear-in-means model excluding education-by-institution cells with more than 300 and 500 students in a cohort, respectively.

[INSERT TABLE 15]

Excluding large cohorts from the estimation has little impact on the estimated peer effects. Even when excluding all education programs where maximum cohort is larger than 300 students in a given year, the estimated peer effect remains at a comparable level and statistically significant. Hence, our results appear not to be driven by measurement error in large peer cohorts.

To further test whether unobserved time-varying education-by-institution characteristics are confounding our results, we perform a series of falsification tests. To do so, we regress dropout on placebo measures of peer ability, namely, the average ability of

[16]Bertoni et al. (n.d.) estimate the effect of peer quality in school on long-term labor market outcomes for Danish adolescents. They find that higher peer quality adversely affects students with high parental education and hypothesize that one of the drivers of this relationship may be ordinal rank effects, which may be larger for high-ability individuals.

previous and future cohorts within the education-by-institution. Since the ability of students is correlated across cohorts, we regress the dropout outcome on lags/leads of cohort ability while controlling for the ability of own peer group. The coefficients on lags/leads of cohort ability capture spillover effects on adjacent cohorts. If unobserved time-varying education-by-institution characteristics are not confounding our estimates, we would expect the spillover effects to diminish as the gap between cohorts expands. Appendix Table A.9 report the coefficient estimates on the average ability of adjacent cohorts.

It is clear from Appendix Table A.9 that the ability of peers have spillover effects on adjacent cohorts. Since delayed students attend courses with younger cohorts, it is not unexpected to find significant impacts from future cohorts on own dropout decision. However, we observe that size and significance of these effects diminishes as the gap increases (forwards and backwards). Hence, these falsification tests provide further evidence that the results are not due to unobserved trends causing a spurious relationship.

8 Conclusion

While an extensive literature on peer effects in higher education exist, the results found are mixed. The evidence is characterized by originating from various more or less selected samples of the general population of higher education students. We have explored the effects of peer quality on the decision to drop out of higher education. Using the universe of higher education enrollments from 1985 to 2010 in Denmark, we have been able to assess the effects of peer quality and how the effects vary in different contexts (field of study, level of program, and the selectivity of the education program). This is an important contribution to the existing literature on peer effects in higher education where studies tend to be based on specific institutions or fields of study. Our findings suggest that the effects of peer quality vary substantially across both field and level of study. Overall, we find that increasing peer quality has favorable effects on dropout behavior. We also document that peer quality has long-term effects on labor market outcomes emphasizing that the effects of peer quality go beyond higher education.

We find significant peer influences in STEM, Social Science, and Arts programs whereas academic ability of peers is less important in Health. By and large, this corresponds to differences found in the peer effects literature. For instance, while several European studies based on data of students in economics or business administration (e.g. Feld and Zölitz, 2017; Booij et al., 2017) estimate significant peer effects, Arcidiacono and Nicholson (2005) find no evidence of peer effects in a sample of US medical students. Moreover, our results suggest only minor influences from peers' quality on the decision to drop out in selective education programs. This result can explain why several studies using data from elite schools find little evidence of peer effects (e.g. Sacerdote, 2001; Zimmerman, 2003; Brodaty and Gurgand, 2016). Finally, we show that peer effects are non-linear and heterogeneous with respect to own ability. In general, our results suggest that high-ability students are less susceptible to the ability composition of their peers. At the same time, high-ability peers have beneficial impacts on the completion rates of lower ability students. In contrast, low-ability peers significantly increase the risk of dropout of similarly low achieving students. In continuation hereof, we show that the definitions of low- and high-ability students have important implications for the estimated peer effects. In particular, we find differences in the peer effect estimates when using a "within-sample" relative rank of students compared to an absolute ranking of students based on the full population. This result can partly explain the mixed findings of the literature with respect to non-linear peer effects.

Arguably some of our findings may reflect the fact that we are considering a somewhat extreme outcome, namely dropping out of higher education. It is perfectly plausible that even for groups where we find no evidence of peer effects in terms of dropout, peer effects may exist in terms of, for example, academic achievement.

Our results highlight that public policies with impacts on student composition have non-negligible effects on dropout rates. Our findings suggest that, on average, a one standard deviation of peers' high school GPA reduces dropout rates by 4.6 percentage points. Students of lower ability are particularly susceptible to the ability composition of their peers, and simultaneously have strong impacts on their peers. Specifically, these findings have important implications for the design of many public policies related to higher education. To a large extent, policy makers can control access to higher education through admission policies. Most admission policies have implications for student composition—for example, admission policies based on a previous test score or GPA—and policy makers can thus benefit from information about how peer quality affects student dropout. In addition, many policy makers are interested in reducing dropout rates, for example through interventions targeted at at-risk students. Our results also provide valuable insights regarding the targeting of such initiatives.

References

Altonji, J. G., Arcidiacono, P. and Maurel, A. (2016), The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects, *in* 'Handbook of the Economics of Education', Vol. 5, Elsevier, pp. 305–396.

Altonji, J. G., Elder, T. E. and Taber, C. R. (2005), 'Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools', *Journal of Political Economy* **113**(1), 151–184.

Ammermueller, A. and Pischke, J.-S. (2009), 'Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study', *Journal of Labor Economics* **27**(3), 315–348.

Angrist, J. D. (2014), 'The Perils of Peer Effects', *Labour Economics* **30**, 98–108.

Arcidiacono, P. (2004), 'Ability Sorting and the Returns to College Major', *Journal of Econometrics* **121**(1-2), 343–375.

Arcidiacono, P., Aucejo, E. M. and Hotz, V. J. (2016), 'University Differences in the Graduation of Minorities in STEM Fields: Evidence from California', *American Economic Review* **106**(3), 525–62.

Arcidiacono, P., Foster, G., Goodpaster, N. and Kinsler, J. (2012), 'Estimating Spillovers using Panel Data, with an Application to the Classroom', *Quantitative Economics* **3**(3), 421–470.

Arcidiacono, P. and Nicholson, S. (2005), 'Peer Effects in Medical School', *Journal of Public Economics* **89**(2-3), 327–350.

Bertoni, M., Brunello, G. and Cappellari, L. (n.d.), 'Who benefits from privileged peers? evidence from siblings in schools', *Journal of Applied Econometrics* **n/a**(n/a).
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.2787*

Black, S. E., Devereux, P. J. and Salvanes, K. G. (2013), 'Under Pressure? The Effect of Peers on Outcomes of Young Adults', *Journal of Labor Economics* **31**(1), 119–153.

Booij, A. S., Leuven, E. and Oosterbeek, H. (2017), 'Ability Peer Effects in University: Evidence from a Randomized Experiment', *The Review of Economic Studies* **84**(2), 547–578.

Bound, J., Lovenheim, M. F. and Turner, S. (2010), 'Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources', *American Economic Journal: Applied Economics* **2**(3), 129–57.
URL: *http://www.aeaweb.org/articles?id=10.1257/app.2.3.129*

Bramoullé, Y., Djebbari, H. and Fortin, B. (2009), 'Identification of Peer Effects through Social Networks', *Journal of Econometrics* **150**(1), 41 – 55.
URL: *http://www.sciencedirect.com/science/article/pii/S0304407609000335*

Brenøe, A. A. and Zölitz, U. (2020), 'Exposure to more female peers widens the gender gap in stem participation', *Journal of Labor Economics* **38**(4), 1009–1054.
URL: *https://doi.org/10.1086/706646*

Brodaty, T. and Gurgand, M. (2016), 'Good Peers or Good Teachers? Evidence from a French University', *Economics of Education Review* **54**, 62–78.

Brunello, G., De Paola, M. and Scoppa, V. (2010), 'Peer Effects in Higher Education: Does the Field of Study Matter?', *Economic Inquiry* **48**(3), 621–634.

Carrell, S. E., Fullerton, R. L. and West, J. E. (2009), 'Does Your Cohort Matter? Measuring Peer Effects in College Achievement', *Journal of Labor Economics* **27**(3), 439–464.

Carrell, S. E., Sacerdote, B. I. and West, J. E. (2013), 'From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation', *Econometrica* **81**(3), 855–882.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W. and Yagan, D. (2011), ' How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star *', *The Quarterly Journal of Economics* **126**(4), 1593–1660.
URL: *https://doi.org/10.1093/qje/qjr041*

Chetty, R., Friedman, J. N. and Rockoff, J. E. (2014), 'Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood', *American Economic Review* **104**(9), 2633–79.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.104.9.2633*

De Giorgi, G., Pellizzari, M. and Redaelli, S. (2010), 'Identification of Social Interactions through Partially Overlapping Peer Groups', *American Economic Journal: Applied Economics* **2**(2), 241–75.

De Paola, M. and Scoppa, V. (2010), 'Peer Group Effects on the Academic Performance of Italian Students', *Applied Economics* **42**(17), 2203–2215.

Epple, D. and Romano, R. E. (2011), Peer Effects in Education: A Survey of the Theory and Evidence, *in* 'Handbook of Social Economics', Vol. 1, Elsevier, pp. 1053–1163.

Feld, J. and Zölitz, U. (2017), 'Understanding Peer Effects: On the Nature, Estimation, and Channels of Peer Effects', *Journal of Labor Economics* **35**(2), 387–428.

Ficano, C. C. (2012), 'Peer Effects in College Academic Outcomes - Gender Matters!', *Economics of Education Review* **31**(6), 1102–1115.

Fischer, S. (2017), 'The Downside of Good Peers: How Classroom Composition Differentially Affects Men's and Women's STEM Persistence', *Labour Economics* **46**, 211–226.

Foster, G. (2006), 'It's Not Your Peers, and It's Not Your Friends: Some Progress toward Understanding the Educational Peer Effect Mechanism', *Journal of Public Economics* **90**(8-9), 1455–1475.

Gemici, A. and Wiswall, M. (2014), 'Evolution of Gender Differences in Post-secondary Human Capital Investments: College Majors', *International Economic Review* **55**(1), 23–56.

Griffith, A. L. and Rask, K. N. (2014), 'Peer Effects in Higher Education: A Look at Heterogeneous Impacts', *Economics of Education Review* **39**, 65–77.

Hoxby, C. (2000), 'Peer Effects in the Classroom: Learning from Gender and Race Variation', *National Bureau of Economic Research Working Paper* **7867**.
URL: *http://www.nber.org/papers/w7867*

Hoxby, C. M. and Weingarth, G. (2005), 'Taking Race Out of the Equation: School Reassignment and the Structure of Peer Effects', *Working paper* .

Humlum, M. K. (2007), 'Estimating the effects of delayed entry into higher education: A discussion', *Nationaløkonomisk Tidsskrift (Danish Economic Journal)* **145**(3), 312–326.

Humlum, M. K., Kristoffersen, J. H. G. and Vejlin, R. (2017), 'College Admissions Decisions, Educational Outcomes, and Family Formation', *Labour Economics* **48**, 215 – 230.
URL: *http://www.sciencedirect.com/science/article/pii/S0927537117302890*

Humlum, M. K., Nandrup, A. B. and Smith, N. (2019), 'Closing or Reproducing the Gender Gap? Parental Transmission, Social Norms and Education Choice', *Journal of Population Economics* **32**(2), 455–500.

Lavy, V. and Schlosser, A. (2011), 'Mechanisms and Impacts of Gender Peer Effects at School', *American Economic Journal: Applied Economics* **3**(2), 1–33.

Lavy, V., Silva, O. and Weinhardt, F. (2012), 'The Good, the Bad, and the Average: Evidence on Ability Peer Effects in Schools', *Journal of Labor Economics* **30**(2), 367–414.
URL: *https://doi.org/10.1086/663592*

Luppino, M. and Sander, R. (2015), 'College Major Peer Effects and Attrition from the Sciences', *IZA Journal of Labor Economics* **4**(1), 4.

Malamud, O. (2011), 'Discovering One's Talent: Learning from Academic Specialization', *ILR Review* **64**(2), 375–405.

Manski, C. F. (1993), 'Identification of Endogenous Social Effects: The Reflection Problem', *The Review of Economic Studies* **60**(3), 531–542.
URL: *http://www.jstor.org/stable/2298123*

Nielsen, H. S., Sørensen, T. and Taber, C. (2010), 'Estimating the Effect of Student Aid on College Enrollment: Evidence from a Government Grant Policy Reform', *American Economic Journal: Economic Policy* **2**(2), 185–215.

OECD (2008), ‘Education at a Glance 2008: OECD Indicators’, OECD Publishing, Paris.

OECD (2009), ‘Highlights from Education at a Glance 2008’, OECD Publishing, Paris.

OECD (2018), *Access to Education, Participation and Progression*, OECD Publishing, chapter Graph B4.2 – Age distribution of first-time entrants into tertiary education (2016).

Sacerdote, B. (2001), ‘Peer Effects with Random Assignment: Results for Dartmouth Roommates’, *The Quarterly Journal of Economics* **116**(2), 681–704.

Sacerdote, B. (2011), Peer Effects in Education: How Might They Work, How Big are They and How Much Do We Know Thus Far?, *in* ‘Handbook of the Economics of Education’, Vol. 3, Elsevier, pp. 249–277.

Sacerdote, B. (2014), ‘Experimental and Quasi-Experimental Analysis of Peer Effects: Two Steps Forward?’, *Annu. Rev. Econ.* **6**(1), 253–272.

Stinebrickner, R. and Stinebrickner, T. R. (2006), ‘What Can be Learned about Peer Effects using College Roommates? Evidence from New Survey Data and Students from Disadvantaged Backgrounds’, *Journal of Public Economics* **90**(8-9), 1435–1454.

Stinebrickner, R. and Stinebrickner, T. R. (2014), ‘A Major in Science? Initial Beliefs and Final Outcomes for College Major and Dropout’, *The Review of Economic Studies* **81**(1), 426–472.
URL: *https://doi.org/10.1093/restud/rdt025*

UFM (2018), Frafald og studieskift. spørgeskemaundersøgelse blandt frafaldne studerende ved videregående udddannelser, Technical report, Uddannelses- og Forskningsministeriet (Ministry of Higher Education and Science), Styrelsen for Forskning og Uddannelse, Copenhagen.

Zimmerman, D. J. (2003), ‘Peer Effects in Academic Outcomes: Evidence from a Natural Experiment’, *Review of Economics and Statistics* **85**(1), 9–23.

9 Tables

Table 1: Descriptive Statistics

	University		College	
	Mean	(Std. Dev.)	Mean	(Std. Dev.)
Personal characteristics				
Female	0.51		0.67	
Age	24.48	(6.07)	25.71	(6.33)
Non-Western Descent	0.04		0.04	
No. siblings	1.38	(1.39)	1.49	(1.60)
Low educated parents	0.06		0.11	
Short educated parents	0.29		0.37	
High educated parents	0.45		0.27	
Parents' income[a]	4.43	(3.74)	3.42	(3.01)
High school degree	0.85		0.68	
Education characteristics				
Average peer ability	0.52	(0.46)	-0.32	(0.38)
Median no. of peers	123		89	
Dropout	0.33		0.26	
First enrollment	0.56		0.79	
No. of observations	696,602		487,488	

This table displays summary statistics separately for university and college enrollments. Average values of individual and education characteristics are reported unless otherwise stated.
[a] In 100,000 DKK. Incomes are deflated using the Danish CPI to 2015 prices.

Table 2: Distribution of Peer Quality

	University	College
Distribution of HS GPA		
Share in decile 1	0.02	0.08
Share in decile 2	0.03	0.08
Share in decile 3	0.04	0.08
Share in decile 4	0.05	0.07
Share in decile 5	0.07	0.07
Share in decile 6	0.09	0.07
Share in decile 7	0.10	0.06
Share in decile 8	0.11	0.05
Share in decile 9	0.13	0.03
Share in decile 10	0.14	0.02
Missing GPA	0.23	0.38

This table shows the distribution of students based on their HS GPA. Decile 1 is the lowest 10th percentile of HS GPA, and decile 10 is the top 10th percentile. Students' ability decile is determined using own high school GPA and the distribution of GPA scores in the year of high school graduation.

Table 3: Decomposition of Variance in Cohort-level Means

	All	University	College
Raw Variation			
Mean HS GPA	0.229	0.506	-0.275
Within variation	0.084	0.091	0.073
Between variation	0.286	0.162	0.119
Total variation	0.371	0.253	0.191
Net of Fixed Effects			
Mean HS GPA	0.003	0.007	-0.006
Within variation	0.076	0.081	0.067
Between variation	0.004	0.004	0.003
Total variation	0.080	0.085	0.070
Net of Fixed Effects and Trends			
Mean HS GPA	-0.000	0.001	-0.003
Within variation	0.060	0.065	0.052
Between variation	0.002	0.002	0.002
Total variation	0.062	0.067	0.054

The table shows a within-between decomposition of the variance in the education-by-institution cohort-level average HS GPA. Following Ammermueller and Pischke (2009), we have weighted the between component by the number of cohorts for the decomposition to add to the total variance in an unbalanced panel.

Table 4: Balancing Tests of the Peer Ability Measure

Dependent Variable:	University Sample	College Sample
Female	0.004	-0.003
	(0.004)	(0.004)
Age	0.038	0.108*
	(0.057)	(0.062)
Birth Month	0.039*	0.055*
	(0.024)	(0.029)
Mothers' Education	0.384	1.060***
	(0.237)	(0.293)
Fathers' Education	0.153	0.907***
	(0.273)	(0.289)
Non-Western Descent	0.002	0.002
	(0.003)	(0.002)
First Education	-0.014	0.006
	(0.014)	(0.006)
Number of Siblings	0.004	-0.001
	(0.006)	(0.008)
Log Parents' Income	-0.001	0.004
	(0.003)	(0.003)
Cohort Size	20.65**	-0.381
	(9.391)	(3.083)
Year FE	✓	✓
Peers' char.	✓	✓
Educ-Inst FE	✓	✓
Educ-Inst trends	✓	✓
Old Peers	✓	✓

Each entry corresponds to a separate regression. The table reports the estimated coefficients on the average HS GPA of peers from regressions of the listed dependent variables. All regressions include education-by-institution FE and trends, year FE, average background characteristics of peers, own HS GPA and year of HS graduation. Standard errors clustered at education-by-institution level are shown in parentheses. The balancing tests have been performed separately for the university and college samples.
$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

Table 5: Linear-in-means Model: The Effects from Peers' Ability on Dropout

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
Peers' HS GPA	-0.011	-0.018**	0.011	-0.067***	-0.046***
	(0.007)	(0.007)	(0.007)	(0.009)	(0.007)
Own HS GPA			-0.057***	-0.056***	-0.056***
			(0.003)	(0.003)	(0.003)
Year FE	✓	✓	✓	✓	✓
Person char.	✓	✓	✓	✓	✓
Peers' char.		✓	✓	✓	✓
Old Peers		✓	✓	✓	✓
Educ-Inst FE				✓	✓
Educ-Inst Trends					✓
N	1,184,090	1,184,090	1,184,090	1,184,090	1,184,090
R^2	0.048	0.059	0.066	0.132	0.156

The table summarizes the results from estimation of the linear-in-means model on dropout using the full sample of higher education enrollments. Each column displays the estimated coefficients on peers' HS GPA and own HS GPA from specifications with increasing number of controls. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table 6: Linear-in-means Model: The Effects from Peers' Ability on Dropout - University Students

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Peers' HS GPA	-0.055***	-0.084***	-0.050**	-0.056***	-0.011
	(0.010)	(0.017)	(0.022)	(0.013)	(0.032)
Own HS GPA	-0.062***	-0.085***	-0.052***	-0.060***	-0.021***
	(0.005)	(0.009)	(0.008)	(0.006)	(0.006)
Difference STEM v. Social Science: p-value = 0.231					
Difference STEM v. Arts: p-value = 0.204					
Difference STEM v. Health: p-value = 0.042					
Difference Social Science v. Arts: p-value = 0.891					
Difference Social Science v. Health: p-value = 0.303					
Difference Arts v. Health: p-value = 0.183					
N	696,602	161,588	295,266	197,446	42,302
R^2	0.177	0.240	0.135	0.170	0.139

The table reports the coefficients on peers' HS GPA and own HS GPA from regressions with dropout as the dependent variable using the university sample. Column (1) shows the results using all university enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses. P-values from Wald tests of differences in the peer effect estimates across field sub-samples are reported.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table 7: Linear-in-means Model: The Effects from Peers' Ability on Dropout - College Students

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Peers' HS GPA	-0.025***	-0.027***	-0.055	-0.042***	-0.003
	(0.008)	(0.013)	(0.042)	(0.015)	(0.010)
Own HS GPA	-0.041***	-0.078***	-0.047***	-0.056***	-0.022***
	(0.003)	(0.006)	(0.011)	(0.008)	(0.003)
Difference STEM v. Social Science: p-value = 0.507					
Difference STEM v. Arts: p-value = 0.433					
Difference STEM v. Health: p-value = 0.157					
Difference Social Science v. Arts: p-value = 0.764					
Difference Social Science v. Health: p-value = 0.222					
Difference Arts v. Health: p-value = 0.030					
N	487,488	87,060	23,310	110,555	266,563
R^2	0.131	0.184	0.230	0.134	0.096

The table reports the coefficients on peers' HS GPA and own HS GPA from regressions with dropout as the dependent variable using the college sample. Column (1) shows the results using all college enrollments. Columns (2)-(5) show the results based on sub-samples of college enrollments in STEM, Social Science, Arts, or Health. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses. P-values from Wald tests of differences in the peer effect estimates across field sub-samples are reported.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table 8: Heterogeneous Peers Effects across Selectivity of Education Program - University Students

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Non-Selective:					
Peers' HS GPA	-0.063***	-0.087***	-0.057***	-0.060***	-0.036
	(0.012)	(0.018)	(0.027)	(0.015)	(0.035)
N	585,053	149,313	241,647	176,447	17,646
Selective:					
Peers' HS GPA	-0.009	-0.024	0.011	-0.042*	-0.051
	(0.015)	(0.029)	(0.033)	(0.024)	(0.030)
N	111,549	12,275	53,619	20,999	24,656
Selective vs. Non-Selective					
Difference:	0.054***	0.063*	0.068	0.018	-0.015
	(0.019)	(0.034)	(0.042)	(0.027)	(0.045)

The table reports the coefficients on peers' HS GPA from regressions with dropout as the dependent variable using the university sample. Column (1) shows the results using all university enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. The top panel displays the results using the sub-sample of non-selective education-by-institutions. The middle panel reports the results from the sub-sample of selective education-by-institutions. The bottom panel reports the differences of the peer effect estimates between selective and non-selective education-by-institutions. An education-by-institution is defined as selective if the cohort-average HS GPA of enrollments is in the top 20th percentile for more than 50% of the years being offered. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1,^{**} p < 0.05,^{***} p < 0.01$.

Table 9: Heterogeneous Peers Effects across Selectivity of Education Program - College Students

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Non-Selective:					
Peers' HS GPA	-0.025***	-0.026*	-0.057	-0.042**	-0.005
	(0.009)	(0.013)	(0.044)	(0.019)	(0.011)
N	423,062	77,375	17,015	86,814	241,858
Selective:					
Peers' HS GPA	-0.017	-0.028	0.045	-0.027	0.011
	(0.014)	(0.039)	(0.034)	(0.019)	(0.019)
N	64,426	9,685	6,295	23,741	24,705
Selective vs. Non-Selective					
Difference:	0.008	-0.002	0.102*	0.015	0.015
	(0.016)	(0.041)	(0.054)	(0.027)	(0.022)

The table reports the coefficients on peers' HS GPA from regressions with dropout as the dependent variable using the college sample. Column (1) shows the results using all college enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. The top panel displays the results using the sub-sample of non-selective education-by-institutions. The middle panel reports the results from the sub-sample of selective education-by-institutions. The bottom panel reports the differences of the peer effect estimates between selective and non-selective education-by-institutions. An education-by-institution is defined as selective if the cohort-average HS GPA of enrollments is in the top 20th percentile for more than 50% of the years being offered. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

Table 10: Heterogeneous Peer Effects across Gender - University Students

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Panel A: All Students					
Peers' HS GPA	-0.059***	-0.091***	-0.053***	-0.058***	-0.014
	(0.012)	(0.019)	(0.025)	(0.014)	(0.027)
Female × peers'	0.010	0.020	0.005	0.003	0.008
HS GPA	(0.009)	(0.017)	(0.016)	(0.013)	(0.028)
Panel B: Female Students					
Female peers' HS GPA	-0.020***	-0.032***	-0.019	-0.024***	0.029
	(0.006)	(0.010)	(0.013)	(0.008)	(0.030)
Male peers' HS GPA	-0.026***	-0.027**	-0.028**	-0.027***	-0.024
	(0.005)	(0.014)	(0.013)	(0.007)	(0.015)
Wald test p-value:	[0.441]	[0.758]	[0.590]	[0.728]	[0.091]
Panel C: Male Students					
Female peers' HS GPA	-0.028***	-0.041***	-0.013	-0.033***	-0.052*
	(0.006)	(0.009)	(0.011)	(0.010)	(0.027)
Male peers' HS GPA	-0.029***	-0.045***	-0.038**	-0.023***	0.008
	(0.007)	(0.014)	(0.015)	(0.008)	(0.021)
Wald test p-value:	[0.903]	[0.807]	[0.084]	[0.376]	[0.151]

The table reports the coefficients on peers' HS GPA from regressions with dropout as the dependent variable using the university sample. Column (1) shows the results using all university enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. Panel A displays the results from a specification where a gender interaction with peers' HS GPA has been included. Panel B reports the coefficients on the gender-specific peer ability measures for the sub-sample of female students. Panel C reports the results for the sub-sample of male students. The Wald tests test the difference between the impacts from male and female peers. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

Table 11: Heterogeneous Peer Effects across Gender - College Students

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Panel A: All Students					
Peers' HS GPA	-0.039***	-0.026**	-0.051	-0.054***	-0.001
	(0.010)	(0.013)	(0.036)	(0.020)	(0.018)
Female × peers'	0.024**	-0.013	-0.015	0.016	-0.002
HS GPA	(0.010)	(0.026)	(0.046)	(0.021)	(0.016)
Panel B: Female Students					
Female peers' HS GPA	-0.040***	-0.015	-0.030	-0.033**	-0.011
	(0.009)	(0.014)	(0.027)	(0.015)	(0.008)
Male peers' HS GPA	0.002	0.008	-0.028	-0.020**	0.005
	(0.004)	(0.024)	(0.036)	(0.008)	(0.003)
Wald test p-value:	[0.000]	[0.426]	[0.954]	[0.413]	[0.068]
Panel C: Male Students					
Female peers' HS GPA	-0.012**	-0.009	-0.015	-0.032*	0.001
	(0.006)	(0.007)	(0.015)	(0.017)	(0.016)
Male peers' HS GPA	-0.011*	-0.013	-0.019	-0.029*	0.005
	(0.006)	(0.009)	(0.029)	(0.015)	(0.006)
Wald test p-value:	[0.924]	[0.736]	[0.904]	[0.906]	[0.785]

The table reports the coefficients on peers' HS GPA from regressions with dropout as the dependent variable using the college sample. Column (1) shows the results using all college enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. Panel A displays the results from a specification where a gender interaction with peers' HS GPA has been included. Panel B reports the coefficients on the gender-specific peer ability measures for the sub-sample of female students. Panel C reports the results for the sub-sample of male students. The Wald tests test the difference between the impacts from male and female peers. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1$,$^{**} p < 0.05$,$^{***} p < 0.01$.

Table 12: Linear-in-means Model: The Effects from Peers' Ability on Labor Market Outcomes

	Dependent Variables: Labor Market Outcomes				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
A. University					
Employment					
Peers' HS GPA	0.006***	0.003	0.003	0.009***	-0.001
	(0.002)	(0.003	(0.005)	(0.003)	(0.007)
Own HS GPA	0.007***	0.007***	0.003***	0.011***	0.003*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
N	538,905	123,492	227,065	158,687	29,661
R^2	0.145	0.134	0.132	0.141	0.185
Labor market earnings					
Peers' HS GPA	9.360***	8.454**	13.535**	11.137***	21.953
	(3.229)	(4.082)	(6.759)	(3.057)	(14.063)
Own HS GPA	22.590***	21.685***	30.165***	15.816***	9.087***
	(1.192)	(1.311)	(2.358)	(0.946)	(1.929)
N	507,983	115,863	213,468	150,925	27,727
R^2	0.290	0.257	0.252	0.213	0.230
B. College					
Employment					
Peers' HS GPA	0.009***	0.005*	0.017	0.012**	0.007***
	(0.002)	(0.003)	(0.011)	(0.005)	(0.002)
Own HS GPA	0.003***	0.008***	0.007***	0.004***	-0.000
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
N	399,397	70,616	17,580	94,626	216,575
R^2	0.089	0.170	0.139	0.108	0.065
Labor market earnings					
Peers' HS GPA	6.890***	3.061	7.874	4.808	4.262**
	(2.020)	(3.589)	(5.391)	(4.351)	(1.812)
Own HS GPA	6.486***	23.503***	19.140***	4.450***	0.207
	(0.997)	(1.471)	(2.902)	(1.300)	(0.761)
N	386,040	66,370	16,971	92,342	210,357
R^2	0.222	0.171	0.299	0.176	0.106

The table reports the coefficients on peers' HS GPA and own HS GPA from regressions with the average employment probability and average labor market earnings 7-12 years after enrollment as the dependent variables. Labor market earnings are measured as 1,000 DKK in 2015 prices. Regressions are estimated based on the population of enrollments in higher education in the years 1985-2006. Panel A (B) reports the results from the university (college) sample. Column (1) shows the results using all enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table 13: Non-linear Peer Effects - University Students

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Panel A: Non-linear Peer Effects					
Fraction of high-GPA peers	-0.073***	-0.105***	-0.138***	-0.063***	0.019
	(0.019)	(0.035)	(0.043)	(0.021)	(0.055)
Fraction of low-GPA peers	0.095***	0.139***	0.002	0.121***	0.109
	(0.034)	(0.051)	(0.066)	(0.044)	(0.088)
Panel B: Heterogeneous Non-linear Peer Effects					
High GPA × Fraction of high-GPA peers	-0.021	-0.037	-0.087**	-0.012	0.029
	(0.020)	(0.044)	(0.044)	(0.025)	(0.058)
High GPA × Fraction of low-GPA peers	-0.056	0.026	-0.200**	0.015	0.110
	(0.052)	(0.089)	(0.096)	(0.069)	(0.128)
Middle GPA × Fraction of high-GPA peers	-0.099***	-0.115***	-0.172***	-0.076***	-0.001
	(0.019)	(0.036)	(0.043)	(0.021)	(0.056)
Middle GPA × Fraction of low-GPA peers	0.079**	0.133***	-0.024	0.104**	0.121
	(0.037)	(0.051)	(0.070)	(0.050)	(0.104)
Low GPA × Fraction of high-GPA peers	-0.119***	-0.122**	-0.179***	-0.150***	0.020
	(0.026)	(0.052)	(0.055)	(0.037)	(0.075)
Low GPA × Fraction of low-GPA peers	0.288***	0.255***	0.182**	0.389***	0.363*
	(0.053)	(0.106)	(0.090)	(0.074)	(0.199)
N	696,602	161,588	295,266	197,446	42,302

The table reports the results from regression with dropout as the dependent variable. Panel A displays the coefficients on the fraction of high-GPA and low-GPA peers. Panel B displays the coefficients on the fraction of high-GPA and low-GPA peers interacted with an indicator of own ability (low, middle, or high). In both panels, the fraction of middle-GPA peers is left as the base. Column (1) shows the results using all university enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Students with a HS GPA below the 20th percentile (of a HS graduation cohort) are defined as low-GPA. Students with a HS GPA above the 80th percentile (of a HS graduation cohort) are defined as high-GPA. Students with a HS GPA between the 20th and 80ht percentile (of a HS graduation cohort) are defined as middle-GPA. The share of high-(low-)GPA peers omits own value. Standard errors clustered at education-by-institution level are reported in parentheses.

$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01.$

Table 14: Peer Effects using Relative Ability Ranking - University Students

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
High GPA × Fraction of high-GPA peers	-0.077**	-0.167***	-0.008	-0.059	-0.015
	(0.032)	(0.056)	(0.076)	(0.047)	(0.117)
High GPA × Fraction of low-GPA peers	0.040	0.040	0.047	0.043	-0.145
	(0.038)	(0.048)	(0.098)	(0.048)	(0.098)
Middle GPA × Fraction of high-GPA peers	-0.124***	-0.268***	-0.081	-0.069	0.068
	(0.032)	(0.065)	(0.071)	(0.044)	(0.095)
Middle GPA × Fraction of low-GPA peers	0.100***	0.161***	0.100	0.139***	-0.111
	(0.038)	(0.048)	(0.103)	(0.048)	(0.089)
Low GPA × Fraction of high-GPA peers	-0.202***	-0.281***	-0.205*	-0.116	-0.158
	(0.046)	(0.105)	(0.080)	(0.073)	(0.095)
Low GPA × Fraction of low-GPA peers	0.173***	0.248***	0.149	0.224***	-0.006
	(0.042)	(0.070)	(0.096)	(0.051)	(0.077)
N	696,602	161,588	295,266	197,446	42,302
R^2	0.178	0.241	0.135	0.170	0.139

The table reports the results from regressions with dropout as the dependent variable. The table displays the coefficients on the fraction of relatively high-GPA and low-GPA peers interacted with an indicator of own relative ability (low, middle, or high). The fraction of middle-GPA peers is left as the base. Column (1) shows the results using all university enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Students with a HS GPA below the 20th percentile (within an education-by-institution) are defined as low-GPA. Students with a HS GPA above the 80th percentile (within an education-by-institution) are defined as high-GPA. Students with a HS GPA between the 20th and 80ht percentile (within an education-by-institution) are defined as middle-GPA. The share of high-(low-)GPA peers omits own value. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

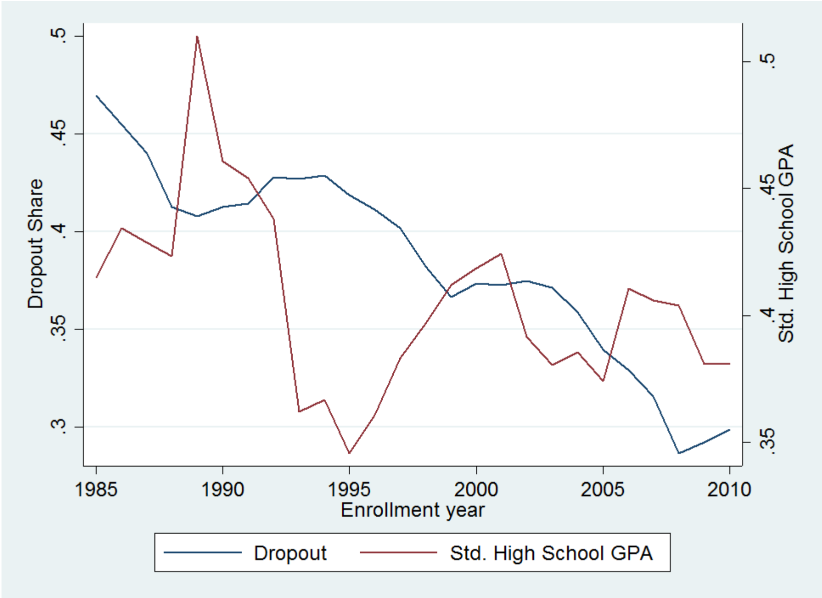
Table 15: Robustness Checks - Exclude Large Cohorts

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	Cohorts <500	Programs max(cohort) <500	Cohorts <300	Programs max(cohort) <300
Panel A: University students					
Peers' HS GPA	-0.055***	-0.060***	-0.053***	-0.051***	-0.046***
	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)
Own HS GPA	-0.062***	-0.060***	-0.061***	-0.059***	-0.059***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
N	696,602	577,933	512,268	519,971	466,073
R^2	0.177	0.193	0.183	0.185	0.181
Panel B: College students					
Peers' HS GPA	-0.025***	-0.025***	-0.025***	-0.025***	-0.021***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Own HS GPA	-0.041***	-0.039***	-0.039***	-0.039***	-0.039***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
N	487,488	482,227	474,617	452,630	409,363
R^2	0.139	0.128	0.129	0.130	0.133

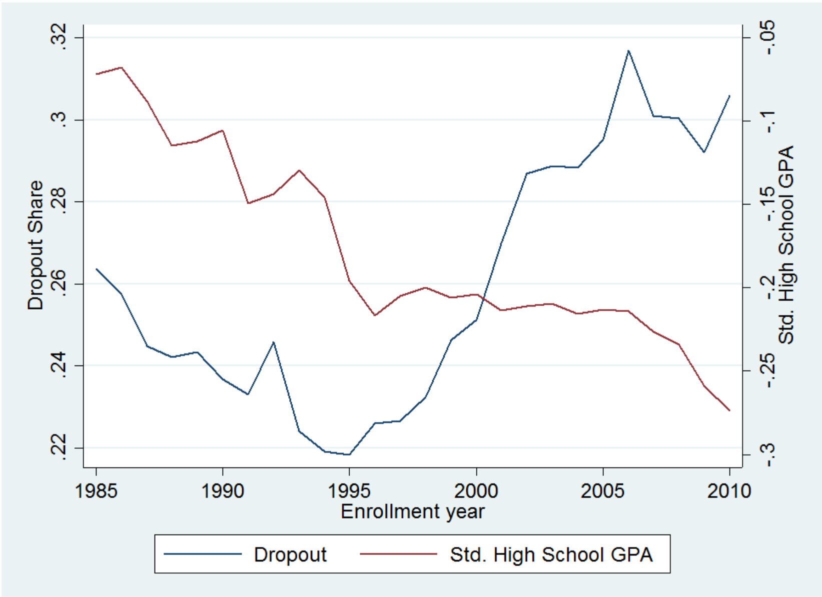
The table reports the coefficients on peers' HS GPA and own HS GPA from regressions with dropout as the dependent variable. Panel A (B) reports the results using the university (college) sample. Column (1) shows the results using all enrollments. Columns (2) and (4) exclude all cohorts with more than 500 and 300 students, respectively, column (3) and (5) exclude all education-by-institution groups where maximum cohort size is greater than 500 and 300, respectively. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

10 Figures

Figure 1: Dropout and Ability Trends



(a) University Enrollments



(b) College Enrollments

The blue lines show the dropout rates from first-time enrollments in higher education programs across years. The red lines plot the average academic ability (measured by the standardized HS GPA) of first-time enrollments in higher education programs.

Appendix Tables

Table A.1: Field of Studies Specification

STEM	Social Science	Arts	Health
45 Natural Sciences 50 Information and Communication Technologies 55 Engineering, engineering trades and manufacturing 57 Mechanics and metal trades 58 Architecture and construction 60 Agriculture, forestry and fisheries 80 Transport services	35 Social Science 40 Business, administration and law 70 Services 75 Security services	20 Education 25 Humanities 27 Audio-visual techniques and media production 30 Arts	65 Health and welfare
Examples of university programs:			
Biology Nanotechnology Statistics Energy technology Physics Geology	Law Psychology Economics Anthropology Sociology Business administration	History Archeology Architecture Philosophy Comparative literature Theology	Physical education Public health science Medicine Odontology Food science
Examples of college programs:			
Software technology Marine and technical engineer Architectural technology and construction Landscape architect Maritime transport and ship management	Journalism Librarian Public Administration Financial economist Media production and management	Computer Graphic Arts Correspondent Fashion design School teacher Actor	Biomedical laboratory sciences Physiotherapy Midwifery Nursing Optician Nutrition and health

Table A.2: Linear-in-means Model: The Effects from Peers' Ability on Dropout - First-Time Enrollment

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Panel A: University students					
Peers' HS GPA	-0.032**	-0.086***	-0.023	-0.020	0.024
	(0.015)	(0.027)	(0.035)	(0.014)	(0.035)
Own HS GPA	-0.088***	-0.113***	-0.079***	-0.079***	-0.034***
	(0.006)	(0.011)	(0.010)	(0.009)	(0.008)
N	393,443	100,548	164,591	102,020	26,284
R^2	0.180	0.248	0.136	0.181	0.140
Panel B: College students					
Peers' HS GPA	-0.023***	-0.025*	-0.053	-0.027*	-0.003
	(0.008)	(0.014)	(0.038)	(0.014)	(0.010)
Own HS GPA	-0.043***	-0.086***	-0.053***	-0.059***	-0.024***
	(0.004)	(0.007)	(0.013)	(0.009)	(0.004)
N	386,498	70,499	16,519	82,690	216,790
R^2	0.136	0.199	0.253	0.136	0.099

The table reports the coefficients on peers' HS GPA and own HS GPA from regressions with dropout as the dependent variable. Panel A (B) reports the results from the university (college) sample. Column (1) shows the results using all enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table A.3: Descriptive Statistics - University Enrollments

	STEM	Social Science	Arts	Health
Personal Characteristics				
Female (0/1)	0.38	0.47	0.65	0.62
Age	23.28	23.90	26.50	23.61
Non-Western Descent (0/1)	0.06	0.04	0.02	0.06
No. siblings	1.40	1.37	1.38	1.49
Low educated parents (0/1)	0.06	0.06	0.06	0.04
Short educated parents (0/1)	0.28	0.31	0.27	0.22
High educated parents (0/1)	0.48	0.45	0.42	0.52
Parents' income[a]	4.56	4.69	3.90	4.64
High school degree (0/1)	0.85	0.87	0.82	0.82
Education Characteristics				
Average peer ability	0.52	0.50	0.45	1.00
Median no. of peers	97	230	68	122
Dropout (0/1)	0.34	0.29	0.42	0.21
First enrollment (0/1)	0.62	0.56	0.52	0.62
Distribution of HS GPA				
Share in decile 1	0.02	0.02	0.02	0.01
Share in decile 2	0.03	0.03	0.03	0.01
Share in decile 3	0.05	0.05	0.05	0.02
Share in decile 4	0.06	0.05	0.05	0.02
Share in decile 5	0.08	0.07	0.07	0.04
Share in decile 6	0.09	0.09	0.09	0.05
Share in decile 7	0.10	0.10	0.09	0.08
Share in decile 8	0.12	0.11	0.11	0.12
Share in decile 9	0.13	0.12	0.12	0.20
Share in decile 10	0.14	0.13	0.11	0.26
Missing GPA	0.19	0.23	0.25	0.20
No. of observations	161,588	295,266	197,446	42,302

The table shows average background characteristics, education characteristics and the distribution of HS GPA of university enrollments. The mean characteristics are reported separately for the sub-samples based on education field.
[a] In 100,000 DKK. Incomes are deflated using the Danish CPI to 2015 prices.

Table A.4: Descriptive Statistics - College Enrollments

	STEM	Social Science	Arts	Health
Personal Characteristics				
Female (0/1)	0.15	0.44	0.69	0.85
Age	24.24	24.43	25.09	26.55
Non-Western Descent (0/1)	0.07	0.04	0.02	0.03
No. siblings	1.48	1.39	1.51	1.51
Low educated parents (0/1)	0.09	0.08	0.11	0.12
Short educated parents (0/1)	0.35	0.35	0.36	0.38
High educated parents (0/1)	0.30	0.37	0.34	0.23
Parents' income[a]	3.52	4.01	3.74	3.20
High school degree (0/1)	0.57	0.84	0.82	0.64
Education Characteristics				
Average peer ability	-0.30	0.04	-0.17	-0.41
Median no. of peers	63	85	187	80
Dropout (0/1)	0.31	0.26	0.32	0.22
First enrollment (%)	0.81	0.71	0.75	0.81
Distribution HS GPA				
Share in decile 1	0.05	0.05	0.07	0.09
Share in decile 2	0.06	0.06	0.08	0.08
Share in decile 3	0.06	0.08	0.10	0.08
Share in decile 4	0.06	0.08	0.09	0.07
Share in decile 5	0.06	0.09	0.09	0.07
Share in decile 6	0.06	0.10	0.09	0.06
Share in decile 7	0.05	0.09	0.08	0.05
Share in decile 8	0.04	0.08	0.06	0.04
Share in decile 9	0.03	0.07	0.05	0.03
Share in decile 10	0.02	0.06	0.03	0.01
Missing GPA	0.50	0.24	0.26	0.41
No. of observations	87,060	23,310	110,555	266,563

The table shows average background characteristics, education characteristics and the distribution of HS GPA of college enrollments. The mean characteristics are reported separately for the sub-samples based on education field.

[a] In 100,000 DKK. Incomes are deflated using the Danish CPI to 2015 prices.

Table A.5: Linear-in-means Model with Course-Specific Skills

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Panel A: University Sample 2000-2010					
Peers' HS GPA	0.013	-0.006	-0.025	0.043**	-0.007
	(0.014)	(0.024)	(0.029)	(0.021)	(0.058)
Peers' HS Math GPA	-0.012	-0.022	-0.016	-0.003	-0.001
	(0.010)	(0.023)	(0.014)	(0.015)	(0.043)
Peers' HS Danish GPA	-0.001	0.014	0.028	-0.064***	0.004
	(0.011)	(0.016)	(0.020)	(0.022)	(0.051)
Own HS GPA	-0.050***	-0.060***	-0.041***	-0.052***	-0.017**
	(0.004)	(0.007)	(0.008)	(0.007)	(0.007)
Own HS Math GPA	-0.021***	-0.043***	-0.021***	-0.007**	-0.021***
	(0.003)	(0.006)	(0.005)	(0.003)	(0.007)
Own HS Danish GPA	0.001	0.004	0.005**	-0.015***	0.011***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
No. of observations	346,994	80,359	154,513	86,471	25,651
Panel B: College Sample 2000-2010					
Peers' HS GPA	0.000	0.020	-0.108	-0.026	0.006
	(0.017)	(0.039)	(0.077)	(0.044)	(0.019)
Peers' HS Math GPA	-0.005	-0.005	-0.153***	-0.061*	0.013
	(0.013)	(0.025)	(0.052)	(0.032)	(0.014)
Peers' HS Danish GPA	0.002	-0.012	0.057	0.105***	0.001
	(0.016)	(0.031)	(0.049)	(0.034)	(0.017)
Own HS GPA	-0.038***	-0.056***	-0.054***	-0.054***	-0.023***
	(0.003)	(0.008)	(0.012)	(0.005)	(0.003)
Own HS Math GPA	-0.011***	-0.034***	0.002	-0.013***	-0.009***
	(0.002)	(0.007)	(0.007)	(0.004)	(0.003)
Own HS Danish GPA	0.004**	0.020***	-0.003	-0.002	0.004*
	(0.003)	(0.008)	(0.012)	(0.005)	(0.003)
No. of observations	223,400	35,592	11,147	47,248	129,413

The table reports the coefficients on peers' HS GPA, peers' math HS GPA, peers' Danish HS GPA, and own values of the variables from regressions with dropout as the dependent variable. Panel A (B) reports the results using university (college) sample. Column (1) shows the results using all enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. All peer ability measures are calculated as the leave-one-out average. Math and Danish HS GPA values are calculated from course-specific grades. Standard errors clustered at education-by-institution level are reported in parentheses. The regressions are performed on the sub-sample of high school graduates in the period 2000-2010 due to limited availability of course-specific grades.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table A.6: Non-linear peer effects - College students

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Panel A:					
Fraction of high-GPA peers	0.014	-0.010	-0.153	0.031	0.049*
	(0.021)	(0.046)	(0.101)	(0.050)	(0.026)
Fraction of low-GPA peers	0.057***	0.039	0.079	0.138***	0.036*
	(0.016)	(0.028)	(0.085)	(0.035)	(0.021)
Panel B:					
High GPA × Fraction of high-GPA peers	0.053*	0.124	-0.023	0.244***	0.028
	(0.030)	(0.086)	(0.102)	(0.078)	(0.038)
High GPA × Fraction of low-GPA peers	0.116***	0.154***	-0.002	0.099*	0.089***
	(0.025)	(0.046)	(0.084)	(0.059)	(0.033)
Middle GPA × Fraction of high-GPA peers	-0.033	-0.052	-0.175*	-0.006	0.008
	(0.024)	(0.061)	(0.101)	(0.053)	(0.032)
Middle GPA × Fraction of low-GPA peers	0.029	0.021	0.050	0.086**	0.005
	(0.018)	(0.028)	(0.096)	(0.039)	(0.023)
Low GPA × Fraction of high-GPA peers	0.179***	0.157**	-0.221**	0.069	0.127***
	(0.041)	(0.071)	(0.100)	(0.094)	(0.046)
Low GPA × Fraction of low-GPA peers	0.040	0.030	0.207**	0.181***	0.054*
	(0.024)	(0.040)	(0.103)	(0.050)	(0.028)
N	487,488	87,060	23,310	110,555	266,563

The table reports the results from regressions with dropout as the dependent variable. Panel A displays the coefficients on the share of high-GPA and low-GPA peers. Panel B displays the coefficients on the share of high-GPA and low-GPA peers interacted with an indicator of own ability (low, middle, or high). In both panels, the share of middle-GPA peers is left as the base. Column (1) shows the results using all college enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Students with a HS GPA below the 20th percentile (of a HS graduation cohort) are defined as low-GPA. Students with a HS GPA above the 80th percentile (of a HS graduation cohort) are defined as high-GPA. Students with a HS GPA between the 20th and 80ht percentile (of a HS graduation cohort) are defined as middle-GPA. The share of high-(low-)GPA peers omits own value. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

Table A.7: Relative Ability Ranking - University Students

	Dependent Variable: Dropout	
	(1) Non-selective	(2) Selective
High GPA × Fraction of high-GPA peers	-0.119***	0.192***
	(0.036)	(0.051)
		[0.000]
High GPA × Fraction of low-GPA peers	0.042	0.002
	(0.043)	(0.064)
		[0.611]
Middle GPA × Fraction of high-GPA peers	-0.176***	0.156***
	(0.036)	(0.051)
		[0.000]
Middle GPA × Fraction of low-GPA peers	0.108**	0.052
	(0.046)	(0.047)
		[0.392]
Low GPA × Fraction of high-GPA peers	-0.241***	0.001
	(0.052)	(0.060)
		[0.002]
Low GPA × Fraction of low-GPA peers	0.209***	0.073
	(0.050)	(0.053)
		[0.060]
N	585,053	111,549
R^2	0.175	0.169

The table reports the results from regressions with dropout as the dependent variable. The table displays the coefficients on the fraction of relatively high-GPA and low-GPA peers interacted with an indicator of own relative ability (low, middle, or high). The share of middle-GPA peers is left as the base. Column (1) shows the results using enrollments in non-selective education-by-institutions. Column (2) shows the results using enrollments in selective education-by-institutions. The square brackets report p-values from Wald tests of the differences in the peer effect estimates between selective and non-selective education-by-institutions. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Students with a HS GPA below the 20th percentile (within an education-by-institution) are defined as low-GPA. Students with a HS GPA above the 80th percentile (within an education-by-institution) are defined as high-GPA. Students with a HS GPA between the 20th and 80ht percentile (within an education-by-institution) are defined as middle-GPA. The share of high-(low-)GPA peers omits own value. Standard errors clustered at education-by-institution level are reported in parentheses.
$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table A.8: Linear-in-means Model: The Effects from Peers' Ability on Dropout - FE without Trends

	Dependent Variable: Dropout				
	(1)	(2)	(3)	(4)	(5)
	All	STEM	Social Science	Arts	Health
Panel A: University students					
Peers' HS GPA	-0.094***	-0.135***	-0.089***	-0.091***	-0.023
	(0.011)	(0.023)	(0.026)	(0.017)	(0.039)
Own HS GPA	-0.071***	-0.094***	-0.064***	-0.069***	-0.029***
	(0.005)	(0.010)	(0.008)	(0.006)	(0.006)
N	696,602	161,588	295,266	197,446	42,302
R^2	0.177	0.240	0.135	0.170	0.139
Panel B: College students					
Peers' HS GPA	-0.036***	-0.057***	-0.171***	-0.025	-0.010
	(0.011)	(0.018)	(0.045)	(0.026)	(0.010)
Own HS GPA	-0.038***	-0.079***	-0.053***	-0.053***	-0.015***
	(0.003)	(0.006)	(0.012)	(0.007)	(0.003)
N	487,488	87,060	23,310	110,555	266,563
R^2	0.131	0.184	0.230	0.134	0.096

The table reports the coefficients on peers' HS GPA and own HS GPA from regressions with dropout as the dependent variable. Panel A (B) reports the results using the university (college) sample. Column (1) shows the results using all enrollments. Columns (2)-(5) show the results based on sub-samples of enrollments in STEM, Social Science, Arts, or Health. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Peers' HS GPA is calculated as the leave-one-out average of all enrollments in an education-by-institution in a given year. Standard errors clustered at education-by-institution level are reported in parentheses.

$^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table A.9: Falsification tests

Outcome: Dropout	Placebo Cohort:									
	Ability of cohort $t-5$	Ability of cohort $t-4$	Ability of cohort $t-3$	Ability of cohort $t-2$	Ability of cohort $t-1$	Ability of cohort $t+1$	Ability of cohort $t+2$	Ability of cohort $t+3$	Ability of cohort $t+4$	Ability of cohort $t+5$
Panel A: University Students										
Average HS GPA	0.016	0.005	0.004	-0.002	-0.022**	-0.042***	-0.029***	-0.014	-0.019*	-0.010
	(0.011)	(0.011)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.012)
N	611,501	633,078	653,381	671,360	686,542	694,567	690,721	684,243	675,433	666,149
Panel B: College Students										
Average HS GPA	-0.013*	-0.019***	-0.012*	-0.024***	-0.020**	-0.025***	-0.022***	-0.002	0.012*	0.009
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.009)	(0.008)	(0.008)	(0.007)	(0.007)
N	429,600	442,465	455,371	468,151	481,001	483,498	474,789	463,037	448,091	431,067

The table reports the results from regressions with dropout as the dependent variable. The table shows the coefficients on the average HS GPA of lagged and lead cohorts within education-by-institutions. Panel A (B) reports the results using the sample of university (college) enrollments. All regressions include education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, average HS GPA of old peers, and HS GPA of own peer group. Standard errors clustered at education-by-institution level in parentheses.
$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix Figures

Figure A.1: Heterogeneous Peer Effects among University Students

The figure plots the estimated coefficients from a regression of dropout on nine measures of the fraction of peers in each HS GPA decile fully interacted with indicators of own ability decile. Each panel plots the interactions with the fraction of peers in a specific decile. Decile 5 (the fraction of peers in the 40th to 50th percentile) is left as the base category. The marginal effects can be interpreted as the impact on dropout from increasing the fraction of peers with HS GPA in a given decile (by decreasing the share of peers with HS GPA in the fifth decile). The regression includes education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Vertical lines indicate the 95 percent confidence interval.

Figure A.2: Heterogeneous Peer Effects among College Students

The figure plots the estimated coefficients from a regression of dropout on nine measures of the fraction of peers in each HS GPA decile fully interacted with indicators of own ability decile. Each panel plots the interactions with the fraction of peers in a specific decile. Decile 5 (the fraction of peers in the 40th to 50th percentile) is left as the base category. The marginal effects can be interpreted as the impact on dropout from increasing the fraction of peers with HS GPA in a given decile (by decreasing the share of peers with HS GPA in the fifth decile). The regression includes education-by-institution FE and trends, year FE, personal background characteristics, peers' average background characteristics, and average HS GPA of old peers. Vertical lines indicate the 95 percent confidence interval.