College as a Marriage Market*

Lars Kirkebøen Edwin Leuven Magne Mogstad

Abstract: What explains the strong educational homogamy and assortativity that we observe among the college educated? We use Norwegian data to address identification and measurement challenges and find that colleges are local marriage markets that matter greatly for whom one marries, not because of the pre-determined traits of the admitted students but as a direct result of attending a particular institution at a given time. The effects of enrollment on homogamy that we uncover are sufficiently large to explain the majority of the strong educational sorting among the college educated in our data. We also examine the mechanisms behind these effects and discuss the implications of our findings for students choosing post secondary education, research on the functioning of the marriage market, and policymakers concerned with inequality and social stratification.

Keywords: assortative mating, college, educational homogamy, field of study, marriage market, search frictions
JEL codes: D13, I23, I24, J12

*Kirkebøen: Research Department, Statistics Norway. E-mail: kir@ssb.no. Leuven: Department of Economics, University of Oslo. Also affiliated with CESifo, IZA, and Statistics Norway. Email: edwin.leuven@econ.uio.no. Mogstad: Department of Economics, University of Chicago; Statistics Norway; NBER. Email: magne.mogstad@gmail.com. We thank Thomas Le Barbanchon, Mikkel Gandil, and seminar participants for valuable feedback and suggestions. The Norwegian Research Council supported this research under projects no. 237840 and 275906.

1 Introduction

There is a long standing literature spanning many disciplines that documents the strong tendency of individuals to sort into internally homogeneous marriages. This sorting has received much attention because the choice of partner has potentially important consequences for inequality within and across generations as well as the reproduction of populations more generally (Schwartz, 2013).

Although education is one of the traits most intensively studied in the assortative mating literature, college graduates are commonly treated as a homogeneous group.[1] An emerging body of descriptive work, however, points to the possibility that the *type* of college education (field or institution) is an important but neglected pathway through which individuals sort into homogeneous marriages. For example, Eika et al. (2019) use Norwegian data to show that while college graduates as a whole were about twice as likely to be married to one another compared to the counterfactual situation where college educated were randomly matched, law graduates are nearly 30 times as likely to be married to one another.[2] Furthermore, Nielsen and Svarer (2009) show that around 20 percent of Danish couples attended the same educational institution.

These descriptive studies raise the question of why college graduates are so likely to marry someone within their own institution or field of study. Several explanations are possible. One is a pure selection story; individuals may match on traits correlated with choice of college field or institution. These traits may be unobserved to the analyst, such as innate ability, tastes or family environment. Another story is one of causation, where the choice of college education causally impacts whether and whom one marries. A causal link can operate through a number of channels, including search frictions or preferences for spousal education.[3]

The goal of this paper is to sort out these explanations and, by doing so, examine the role of colleges as marriage markets. Distinguishing between selection and causation as explanations for the observed educational homogamy and assortativity may be important for several agents in the economy: students

[1] See literature reviews in Blossfeld (2009), Han and Qian (2020), and Eika et al. (2019).

[2] See also Bičáková and Jurajda (2016) and Han and Qian (2020), who show strong assortative mating by post-secondary field of study in various OECD countries and in the US, respectively.

[3] Both theory and evidence suggest marriage decisions are increasingly driven by returns to matching on similarities (e.g. due to leisure complementarities), rather than potential gains from trade (see the review in Juhn and McCue (2017)).

choosing post-secondary education; researchers studying the functioning of the marriage market; and policymakers concerned with inequality and social stratification. Yet, there is little if any research on educational homogamy and sorting that makes this distinction.

The context of our study is Norway's post-secondary education system. Our work draws on two strengths of this environment. First, Norwegian register data allow us to observe not only people's choice of college education (institution and field) and workplace, but also if and who they marry (or cohabit with). Second, a centralized admission process creates instruments for choice of college eduction from discontinuities that effectively randomize applicants near unpredictable admission cutoffs into different institutions and fields of study.[4] Thus, differences in marriage market outcomes across these applicants are due to the institution or field to which they are exogenously assigned, as opposed to their pre-existing traits.

Our instrumental variables estimates are summarized with six broad conclusions. First, the type of post-secondary education is empirically important in explaining whom but not whether one marries. Indeed, the magnitude of the effects of enrollment on homogamy are sufficiently large to explain a majority of the strong educational homogamy and assortativity that we observe among the college educated in our data. Second, educational homogamy is economically important as the potential earnings of partners in homogamous matches are materially different from the potential earnings of partners in non-homogamous matches. This remains true even after we account for selection into type of education, which suggests that non-homogamous partners are far from perfect substitutes to homogamous partners in terms of potential earnings. Third, enrolling in a particular institution makes it much more likely to marry someone from that institution. These effects are especially large if individuals overlapped in college, are sizable even for those who studied a different field and are not driven by geography. Fourth, enrolling in a particular field increases the chances of marrying someone within the field but only insofar the individuals attended the same institution. Enrolling in a field makes it no more likely to marry someone from other institutions with the same field. Fifth, the effects of enrollment on educational homogamy and assortativity vary systematically

[4] Kirkeboen et al. (2016) use these discontinuities to show that earnings payoffs vary a lot by post-secondary field of study and less by institution. The results highlight the limitations of treating college educated as a homogeneous group.

across fields and institutions, and tend to be larger in more selective and higher paying fields and institutions. Sixth, only a small part of the effect of enrollment on educational homogamy can be attributed to matches within the same workplace. Lastly, the effects on the probability of marrying someone within their institution and field vary systematically with cohort-to-cohort variation in sex ratios within institutions and fields. As discussed in greater detail later, this finding is at odds with the assumption in canonical matching models of large and frictionless marriage markets.[5]

Taken together, our findings suggests that colleges are effectively local marriage markets, mattering greatly for whom one marries, not because of the predetermined traits of the students that are admitted but as a direct result of attending a particular institution at a given time. Our ability to credibly distinguish between causal effects of educational choices and selection makes these findings relevant for both students, researchers, and policymakers. If the strong assortative mating by institution and field simply reflected selection, then it would (or at least should) be immaterial for individuals' incentives and returns to educational choices; even in the absence of choosing a particular field or institution, they would have been just as likely to match with the same partners. The ability so separate causal effects from selection is also essential for researchers studying the functioning of the marriage market. Otherwise, it would be impossible to tell whether the strong sorting that is observed by field and especially institution would have happened even in the absence of individuals' educational choices.

Policymakers concerned with inequality and social stratification may also be interested in our finding that the post-secondary system contributes to increased social stratification in the form of educational homogamy. We are able to reach this conclusion precisely because we are able to separate causal effects from selection. The fact that the causal effects on homogamy are much stronger when there are similar number of men and women within an institutions and field suggests a trade-off or tension between the concern about homogamy and gender equality in the higher education system. An underappreciated (and possible unintended) consequence of fields and institutions becoming more similar in their representation of men and women is that educational homogamy should increase. Thus, equality in one dimension creates inequality in another

[5] Chiappori (2020) reviews theory and empirics of the marriage market, discussing both frictionless matching models and search models with frictions.

dimension.

Our paper contributes to a growing literature that documents educational homogamy and assortativity and tries to understand their causes and consequences. The closest studies to our work are concerned with drawing causal inference about how the choice of post-secondary education affects the quality of an individual's spouse. Kaufmann et al. (2013) study admission to elite higher education institutions in Chile. Applying a regression discontinuity design to the admission system, they find that attending a higher ranked university has a sizable effect on the quality of an individual's spouse. Artmann et al. (2018) use admission lotteries for four oversubscribed programs in the Netherlands, and find that field of study matters for partner choice. Neither of these studies document how and why the choices of post-secondary education affect educational homogamy and assortativity.

A related body of work seeks to quantify the importance of search frictions and meeting opportunities for assortative mating. One of these is Nielsen and Svarer (2009), who use Danish data to document the extent to which individuals match on education length and type.[6] They find that around half of the systematic sorting on education can be attributed to the tendency of individuals to marry someone who went to the same educational institution or to an institution nearby. Nielsen and Svarer conclude that this may be due to search frictions or selection of people with the same preferences into the same institution. To address this identification challenge, a few studies have taken advantage of detailed data from the dating market. Hitsch et al. (2010) find that search frictions may play an important role in explaining the observed matching pattern by education at an online dating site. Belot and Francesconi (2013) use data from a speed dating agency to identify the role of opportunities separately from that of preferences. Their findings suggest the role of individual preferences is outplayed by that of opportunities.

We complement the existing work in several ways. First, we show that the choices of post-secondary education are empirically important in explaining whom but not whether one marries while addressing concerns about self-

[6] See also Mansour and McKinnish (2014), who use survey data from the US to document that same-occupation matching is strongly related to the sex composition of the occupation. To distinguish between a preferences explanation and a search cost explanation, they investigate whether women accept lower-wage husbands if they match within-occupation compared to if they do not, and how this wage gap varies with the sex composition of the occupation.

selection into education based on unobservables. Second, the admission system we study creates exogenous variation in both field and institution choice, which allows us to disentangle the relative importance of institution and field of study. Third, because we can follow individuals through the education system and into the labor market, we can jointly examine the choice of education and workplace for educational homogamy. Lastly, given our detailed and large population panel data, we are able to estimate market-specific matching functions that helps investigate the impact of local market tightness and size on educational homogamy.

The remainder of the paper proceeds as follows. Section 2 describes the secondary education sector in Norway and its admission process. Section 3 presents the data and sample restrictions. Section 4 defines and describes educational homogamy and assortativity among high educated in Norway. Section 5 provides a graphical depiction of the relevance and validity of our research design, while Section 6 turns to the formal econometric model. Section 7 presents our main findings on how the choices of post-secondary institution and field affect educational homogamy and assortativity, while Section 8 examines the ways in which this educational homogamy and assortativity arise. The final section concludes.

2 Institutions, admission process, and identification strategy

In this section, we describe the secondary education sector in Norway and its admission process, laying the groundwork for what we do in the empirical analysis.

2.1 Post-secondary education sector

During the period we study, the Norwegian post-secondary education sector consisted of a handful of public universities and a number of public and private university colleges. The vast majority of students attend a public institution, and even the private institutions are funded and regulated by the Ministry of Education and Research. A post-secondary degree normally lasts 3–5 years. The universities all offer a wide selection of fields. By comparison, the university colleges rarely offer fields like Law, Medicine, Science, or Technology, but tend to offer professional degrees in fields like Engineering, Health, Business,

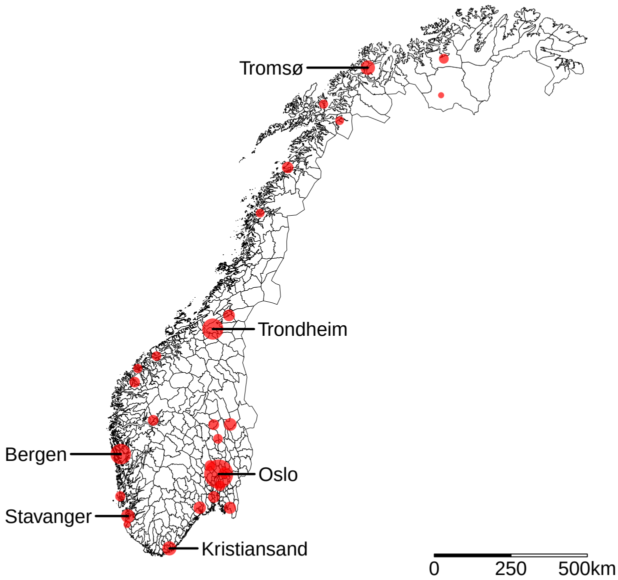


Figure 1. Higher education enrollment

and Teaching. There are generally no tuition fees for attending post-secondary education in Norway, and most students are eligible for financial support (part loan/part grant) from the Norwegian State Educational Loan Fund.

The main universities are located in the major cities of each of the five regions: Bergen and Stavanger (West), Oslo (East), Kristiansand (South), Trondheim (Central) and Tromsø (North). In addition, there are a few other universities and several university colleges spread across the country. Figure 1 displays the distribution of the post secondary student population across Norwegian municipalities in the years 1998–2004.[7] As can be seen in the Figure, in most (393 of the 422) municipalities there are no colleges or universities and, thus, few if any students. The large majority of students (about 60 percent) live in Oslo, Bergen or Trondheim, the three biggest cities of Norway. There is also a sizable student population in a few other municipalities, such as Tromsø, Kristiansand and Stavanger.

[7]There are generally long driving distances between (populated areas of) the different local labor markets, which are mostly far apart or partitioned by mountains or the fjord-broken shoreline. Thus, students attending a given university/university college typically live in the same commuting zone. Most students live off campus, either with parents or in apartment rentals.

2.2 Admission process

The admission process to post-secondary education is centralized. Individuals submit their application to a single organization, the Norwegian Universities and Colleges Admission Service, which handles the admission process to all universities and most university colleges. The unit in the application process (program) is the combination of field and institution (e.g. Teaching at the University of Oslo).

Every year in the late fall, the Ministry of Education and Research decides on funding to each field at every institution, which effectively determines the supply of slots. While some slots in some programs are reserved for special quotas (e.g. students from northernmost part of Norway), the bulk of the slots are for the main pool of applicants. For many programs, demand exceeds supply. Programs for which there is excess demand are filled based on an application score derived from high school GPA. Individual course grades at high school range from 1 to 6 (only integer values), and GPA is calculated as 10 times the average grade (up to two decimal places). A few extra points on the application score are awarded for choosing specific subjects in high school. For some programs, the application score can also be adjusted based on ad-hoc field specific conditions unrelated to academic requirements (e.g. two extra points for women at some male-dominated fields). Additionally, applicants can get some compensation in their application score depending on their age, previous education and fulfillment of military service.

On applying, applicants rank up to fifteen programs. Information about what fields are offered by the different institutions is made available in a booklet that is distributed at high schools. The deadline for applying to programs is mid-April. This is the applicants' first submission of program rankings. They can adjust their rankings until July. New programs cannot be added, but programs can be dropped from the ranking. Once the rankings are final in July, offers are made according to a sequential dictatorship mechanism where the order is determined by the applicants' application score: the highest ranked applicant receives an offer for her preferred program; the second highest applicant receives an offer for her highest ranked program among the remaining programs; and so on. This is repeated until either slots run out, or applicants run out.

This procedure generates a first set of offers which are sent out to the applicants in late July. Applicants then have a week to accept the offer, if they get one. Irrespective of whether they accept, applicants can choose to remain on a waiting list for preferred program options, or withdraw from the application process. The slots that remain after the first round are then allocated in a second round of offers in early August among the remaining applicants on the waiting list. These new offers are generated following the same sequential dictatorship mechanism as in the first round. Since applicants in this second round can only move up in the offer sequence, second round offers will either correspond to first round offers, or be an offer for a higher ranked program. By choosing to remain on a waiting list the applicants accept that their first round offer is automatically discarded if they get a higher-ranked offer in the second round. In mid-August, the applicants begin their study in the accepted field and institution. If students want to change field or institution, they usually need to participate in next year's admission process on equal terms with other applicants.

2.3 Admission thresholds and identification strategy

As described above, the admission process to post-secondary education generates a setup where applicants scoring above a certain threshold are much more likely to receive an offer for a program they prefer as compared to applicants with the same program preferences but marginally lower application score. As illustrated in Table 1, this process creates discontinuities which effectively randomize applicants near unpredictable admission cutoffs into different programs, fields and institutions.

We first consider how to use these discontinuities to identify the impact of choosing one type of program as compared to another. To this end, panel (a) of Table 1 is sufficient. This panel presents an example of an application where the applicant is on the margin of getting different field offers from the same institution. Suppose the applicant has an application score of 49. In this case, she would receive an offer for her 3rd ranked program. This defines her *preferred program* in the local program ranking around her application score, namely the program consisting of field 2 at institution A. We can now compare her to an applicant with the same ranking of programs, but who has a slightly lower application score of 47. This applicant has the same programs in the local ranking around her application score. However, because of the marginally lower score, she does not receive an offer of the preferred program, field 2 at institution A.

Table 1. Illustration of comparisons used to identify threshold crossing effects of programs, fields and institutions

(a) Fields

Program Ranking	Inst.	Field	Cutoff
1st best	A	1	57
2nd best	B	1	52
3rd best	A	2	48
4th best	A	3	45

	Application score = 49		
Local Ranking	Inst.	Field	Offer
Preferred	A	2	Yes
Next-best	A	3	No

	Application score = 47		
Local Ranking	Inst.	Field	Offer
Preferred	A	2	No
Next-best	A	3	Yes

(b) Institutions

Program Ranking	Inst.	Field	Cutoff
1st best	B	1	52
2nd best	A	2	48
3rd best	B	2	46
4th best	B	3	43

	Application score = 49		
Local Ranking	Inst.	Field	Offer
Preferred	A	2	Yes
Next-best	B	2	No

	Application score = 47		
Local Ranking	Inst.	Field	Offer
Preferred	A	2	No
Next-best	B	2	Yes

By comparing the outcomes of applicants like these – with the same preferred program and application scores just below and above the program's admission cutoff – we can estimate the effect of crossing the admission threshold to the preferred program. As long as individuals are not able to perfectly sort around the cutoffs, we can rule out that differences in their outcomes are driven by unobserved heterogeneity in preferences, ability and other confounders. In Section 4, we will report such threshold crossing effects for a wide range of outcomes. Next, we will, in Section 5, use threshold crossing as an instrument for enrollment in a particular program. This instrumental variables approach allows us to draw inferences about how the choice of college education affects whether and whom one marries.

It is important to observe that the admission process creates exogenous variation in not only programs but also fields and institutions. This allows us to quantify the relative importance of field versus institution of study for marriage and educational homogamy. To see this, it is useful to consider both panels of Table 1. In panel (a), the two applicants are on the margin of getting an offer for the same field but from different institutions. One applicant has a application score of 49 and, therefore, receives an offer from her *preferred field* in the local field ranking around her application score, namely field 2. The other applicant has the same ranking of fields and institutions, but is not offered field 2 because she has a slightly lower application score of 47. By comparing the outcomes of these applicants we can estimate the effect of getting an offer from the preferred field 2. Panel (b) gives another example where two applicants are on the margin of getting an offer for the same field but from different institutions. One applicant has a application score of 49. Thus, she receives an offer from institution A, her *preferred institution* in the local institution ranking around her application score. The other applicant has the same ranking of institution, but is not offered institution A because she has a slightly lower application score of 47. By comparing the outcomes of these applicants we can estimate the effect of getting an offer from the preferred institution. In addition to estimating the causal effects of being offered particular fields or institutions, we can also use threshold crossing indicators as instruments to infer the consequences of enrolling in these fields or institutions.

Finally note that applicants can be on two margins. For example, the applicant with a score of 49 in (a) can be on the margin between (B, 1) and (A, 2) or between (A, 2) and (A, 3). In our analysis below we use both margins. However, only about 15–20 percent of applicants are observed on two margins, and our estimates do not materially change if we exclude these applicants (see robustness analyses in Section 7).

3 Data and descriptive statistics

3.1 Data sources, sample selection, and definition of key variables

Our analysis employs several data sources from Norway that we can link through unique identifiers for each individual, spouse and parent. We start with the application records from the Norwegian Universities and Colleges Admission Service. These records give information on nearly all applications to post-secondary education in Norway for the years 1998 to 2004. We select the application cohorts 1998–2004, where 1998 is the first cohort for which data is available. Stopping at 2004 allows us to study marriages and cohabitation in a balanced panel formed within 13 years after application.[8] We retain the individuals' first observed application, also requiring that they – at the time of application – have no post-secondary degree, are younger than 27 years, and are neither married nor cohabiting. We also drop applicants who have an application score further than 2 standard deviations away from the threshold, and who have missing information on completed education 13 years after applying.

The application records provide information for each applicant on his or her ranking of programs, application scores, offers received and enrollment decisions. In addition, we observe the second (and final central) round admission cutoffs (if any) for each program in every year. We merge these records with administrative registers provided by Statistics Norway that cover every resident from 1967 to 2017. For each year, it contains individual socioeconomic information (including sex, age, marital and cohabiting status, educational attainment, and earnings). Following Norwegian official statistics, we define a *match* as a man and a woman who are living at the same address and are either married or cohabitants, or have a child together. The data allow us to construct measures of the educational attainment, earnings and socio-economic background of both the applicants and their partners (if any). Partners are all partners in the population, not restricted to applicants/enrollees, and partner's education is completed education.

The information on educational attainment includes both the completed field and the institution from which individuals graduate. The background variables are based on information about parental education (both for the mother and father), income of the father, and the immigrant status of the family. This information is pre-determined in the context of our analysis, and refers to the year when the applicant was 16 (fathers' earnings are averages at ages 16 and 19).

In our main analysis, we consider an estimation sample of 110,345 applicants who apply for at least two programs, where the most preferred program needs to have an admission cutoff, and the next-best alternative must have a lower cutoff (or no binding cutoff). This ensures that we have information on the preferred program, and a source of identification (potentially binding admission cutoffs) in our analysis. In the analysis of field and institution of study, we construct estimation samples in a similar fashion.

[8] The panel is nearly perfectly balanced. Only a small number of individuals (about 2 percent) drop out at some point during the 13 year period mainly due to emigration.

Table 2. Summary statistics of applications: Rankings and offers

	Sample Mean	St. Deviation
Rankings		
# Programs ranked	6.8	4.1
# Fields ranked	3.5	2.1
# Institutions ranked	4.0	2.7
Offers		
Rank of best offer	2.6	2.5
Offered 1st rank	0.47	
Offered 2nd rank	0.24	
Offered 3rd rank	0.12	
No offer	0.10	

Note: This table presents summary statistics of the main estimation sample, consisting of 110,345 applicants.

To construct measures of assortativity, we also need to take a stand of the population of potential partners. About 70 percent of all applicants in our sample are 19–21 years in the application year. We therefore approximate the population of potential partners by other college graduates who were also aged 19–21 in the year of application. Empirically, the results are highly robust to the exact choice of this age range. The reason is that the measures of assortativity we use are a function of the shares of same-sex and opposite-sex graduates with the same education which vary relatively little from cohort to cohort in our sample period.

3.2 Summary statistics of applicants

Table 2 reports summary statistics for the applications of our estimation sample. These applicants listed, on average, about 7 programs, across 3–4 different fields and institutions. While a substantial fraction is offered their first ranked program, the average offer is for the 3rd ranked program. 10 percent are not offered any program.

Table 3 shows descriptive statistics for key characteristics and outcomes of our estimation sample. The majority of applicants, about 62 percent, is female. The applicants are, on average, between 20 and 21 years old when we observe

Table 3. Summary statistics of applicants: Key characteristics and outcomes

	Sample Means	
Pre-determined characteristics:		
- Age	20.7	
- Female	0.62	
- Immigrant	0.04	
- College educated parents	0.50	
- Application score	49.0	
	Application year	Within 13 years
Enrollment in:		
- Any college	0.77	0.97
- Preferred program	0.33	0.44
- Preferred field	0.36	0.55
- Preferred institution	0.42	0.56
Marriage:		
- Any spouse		0.80
- College-educated spouse		0.50

Note: This table presents summary statistics of the main estimation sample, consisting of 110,345 applicants.

them applying for the first time.[9] About 50 percent of the applicants has a high-educated mother or father. Only 4 percent of the applicants are immigrants. Table 3 shows that 77 percent of the applicants enroll in some higher education in the application year, 36 percent enroll in their preferred field and 42 in their preferred institution, and 33 percent enroll in their preferred program (i.e., combination of field and program). Within 13 years after applying, 97 precent of applicants have enrolled in some higher education, 44 percent in their preferred program at the time of application, and about 55 percent in their preferred field or institution. Examining the probability of matching within 13 years after applying, we find that 80 percent of the applicants are matched with a spouse or partner while 50 percent have a college-educated spouse. As shown in Appendix Figure A1, these marriage patterns are fairly similar across cohorts.

[9]In Norway, students graduate from high school in the year they turn 19, after which many serve in the military, travel, or work for a year or two before enrolling in post-secondary education.

Table 4. Summary statistics: Educational homogamy and assortativity

	Homogamy			Assortativity	
	Observed h	Random h^r	Maximal h^m	Absolute $h - h^r$	Rescaled $\frac{h-h^r}{h^m-h^r}$
Homogamy:					
– Program	0.046	0.002	0.545	0.044	0.082
– Field	0.062	0.013	0.578	0.049	0.087
– Institution	0.121	0.012	0.705	0.109	0.157

Note: This table presents summary statistics of the main estimation sample, consisting of 110,345 applicants. Observed homogamy is the share of applicants and partners with same completed program, field or institution. Random homogamy is constructed as the share matched times the share of potential partners with the relevant type of education for each type of education, sex and application year. Maximal homogamy is the share of students that could potentially be matched to an opposite-sex partner with the same type of education in the population of potential partners. Rescaled assortativity is constructed by rescaling average absolute assortativity with the average difference between maximal and random homogamy.

4 Educational homogamy and assortativity

In Table 4, we dig deeper into the matching patterns in the estimation sample. The first column reports *educational homogamy* rates, as measured by the share of the applicants whose first observed match is to someone with the same type (program, field or institution) of college education. The program homogamy rate is about 4.6 percent. By comparison, 6.2 percent are homogamous with respect to field. This means that 1 out of 8 of the college educated couples have degrees in the same field. The institution homogamy rate is as large as 12 percent, which means that both spouses have graduated from the same institution in 1 out of 4 of the college educated couples.

The fact that the homogamy rate is largest at the level of the institution does not necessarily imply that assortativity is stronger by institution as compared to field or program. This is because homogamy not only depends on the degree of sorting but also on the number of men and women with each type of education. Homogamy rates could therefore be larger for institutions and fields than for programs even if people would have been matched randomly. The second results column of Table 4 confirms that this is indeed the case in our data. If we randomly match applicants and partners with potential partners, then the homogamy rate is close to zero for programs and around 1.5 percent for institutions and fields. This shows that individuals are much more likely to match with someone from the same institution, field and program as compared to what one would observe under random matching.

To quantify the amount of sorting, it is useful to supplement estimates of educational homogamy with measures of assortativity, which is regularly defined as a mating pattern in which individuals with similar traits mate with one another more frequently than would be expected under a random mating pattern. This definition suggests that *educational assortativity* can be measured by comparing the observed homogamy rates to those produced by random matching of men and women. In the fourth results column of Table 4, we compute this measure of assortativity for programs, fields and institutions. These numbers show that applicants with the same type of education match much more frequently than what would be expected under a marriage pattern that is random in terms of education. This positive assortativity occurs for programs, fields and institutions. However, the degree of assortativity is heterogeneous and varies depending on the margin of college education one considers. The applicants were 10 percentage points more likely to be married to someone with a degree from the same institution compared to random mating; for program or field, they were 4 and 5 percentage points more likely to match than expected under random mating patterns.

When interpreting measures of educational assortativity it is important to observe that the values they can take are constrained by the marginal distributions of education among men and women. This is most easily illustrated with two types of education. If men are women have equal distributions of education then perfect assortativity is feasible (i.e., the homogamy rate can be one). If all men have education of one type while all women have education of the other type, then assortativity is infeasible (i.e., the homogamy must be zero). We follow Liu and Lu (2006) and recenter homogamy rates relative to random matching and scale it relative to maximum feasible homogamy as follows

$$R = \frac{h - h^r}{h^m - h^r} \tag{1}$$

where h is the observed homogamy rate, h^r is the homogamy rate we would observe under random matching, and h^m is the maximal attainable homogamy. Using this rescaled measure the absence of assortativity corresponds to $R =$

0, and the maximum positive level of assortativity that is attainable given the distribution of education on the two sides of the market is denoted by $R = 1$. For a given education $E = e$ (program, field, or institution), let s_e^s denote the share of same-sex graduates and s_e^o the share of opposite-sex graduates. Then:

$$h_e^r = s_e^s \times s_e^o \tag{2}$$

and

$$h_e^m = \frac{\min(s_e^s, s_e^o)}{s_e^s}. \tag{3}$$

The overall random and maximum homogamy rates (reported in columns 3 and 4 of Table 4) follow from averaging h_e^r and h_e^m with education shares $s_e \equiv \Pr(E = e)$.

The final column in Table 4 reports the resulting rescaled homogamy rates. These centered and rescaled measures show that assortativity is high at all education margins. For programs and fields of study around 8 percent of maximum feasible assortativity is realized, and this number is twice as high at 16 percent at the institution level.

5 The effects of threshold crossing

Before we analyze how the choice of college education affects whether and whom one marries, it is useful to understand what does and does not change at the admission thresholds. We start with documenting that there is no evidence of strategic sorting around the thresholds and little if any impact of threshold crossing on the likelihood of enrolling or ultimately completing college. What does change as a result of crossing the admission thresholds is the type of education that one enrolls in and completes, as well as whom one marries. Threshold crossing does not, however, affect the overall likelihood of matching.

To estimate the effect of threshold crossing we first standardize and center individuals' application score at the threshold. We then estimate local-linear regressions on both sides of the cutoff and use this to estimate the average treatment effect at the application cutoff. We use the implementation of Calonico et al. (2014), a common MSE-optimal bandwidth selector and report the resulting estimate and bias-corrected confidence interval (b-c CI) in the figures below. We also estimate OLS regressions with a linear spline in the standardized application score and report the coefficient on the threshold crossing dummy. The local-linear regression estimates support the parametric specification of the linear-spline regressions, which deliver estimates that are always in line with the non-parametric ones and also are more precise.

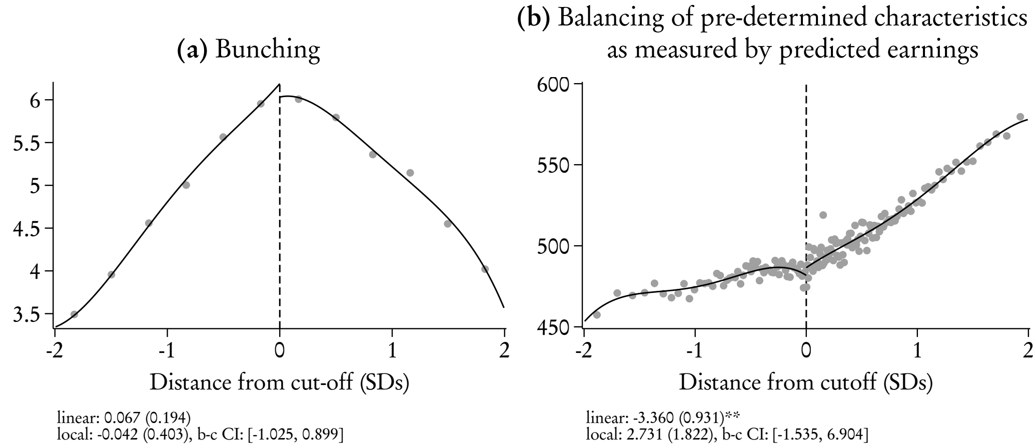
Assessing the validity of the regression-discontinuity design

A potential threat to the regression-discontinuity design that we are exploiting here is that people might try to sort themselves above the cutoff in order to receive an offer for their preferred program of study. If such sorting occurs we would expect to observe discontinuities around the cutoffs in the density of applicants and in their observed characteristics. We investigate this in turn.

Figure 2 pools all the programs and admission cutoffs. The data is normalized so that zero on the x-axis represents the admission cutoff to the preferred program, and observations to the left (right) of this cutoff have therefore an application score that is lower (higher) than the cutoff. We plot the unrestricted means in bins and include regression lines on each side of the admission cutoff. What matters for our research design is that there is no discontinuous jump in probability mass at zero, since that would point to sorting. As can be seen in Figure 2, there is no indication that applicants are able to strategically position themselves around the application boundary, and the test proposed by McCrary (2008) does not reject the null hypothesis of no sorting at conventional levels.

A complementary approach to assess the validity of the research design is to investigate covariate balance around the cutoffs. We consider several individual characteristics that correlate with earnings: gender, cohort, application score, parental education and immigrant status. We construct a composite index of these pre-determined characteristics, namely predicted earnings, using the coefficients from an OLS regression of earnings on these variables. Figure 2 shows average predicted earnings in small intervals on both sides of the pooled application cutoffs and (global and local) linear regression lines. There is no indication that applicants are materially different in terms of observables at the application boundaries.

Taken together, the results in Figure 2 suggest that students do not sort themselves around the admission cutoffs. The absence of sorting around the cutoffs is consistent with key features of the admission process. First, the exact admission cutoffs are unknown both when individuals do their high school exams



Note: Panel (a) shows log density of applicants. Panel (b) shows an index of pre-determined characteristics, constructed as predicted earnings. Estimates in notes are with global linear splines and local linear regressions using a triangular kernel using Calonico et al. (2014). Standard errors in parentheses. Bias-corrected confidence intervals from Calonico et al. (2014) in brackets. Graphs are constructed using the `rdplot` command from Calonico et al. (2014). Dots represent bin means, bins are selected with the IMSE-optimal quantile-based method using polynomial regression. The lines represent fourth-degree global splines.

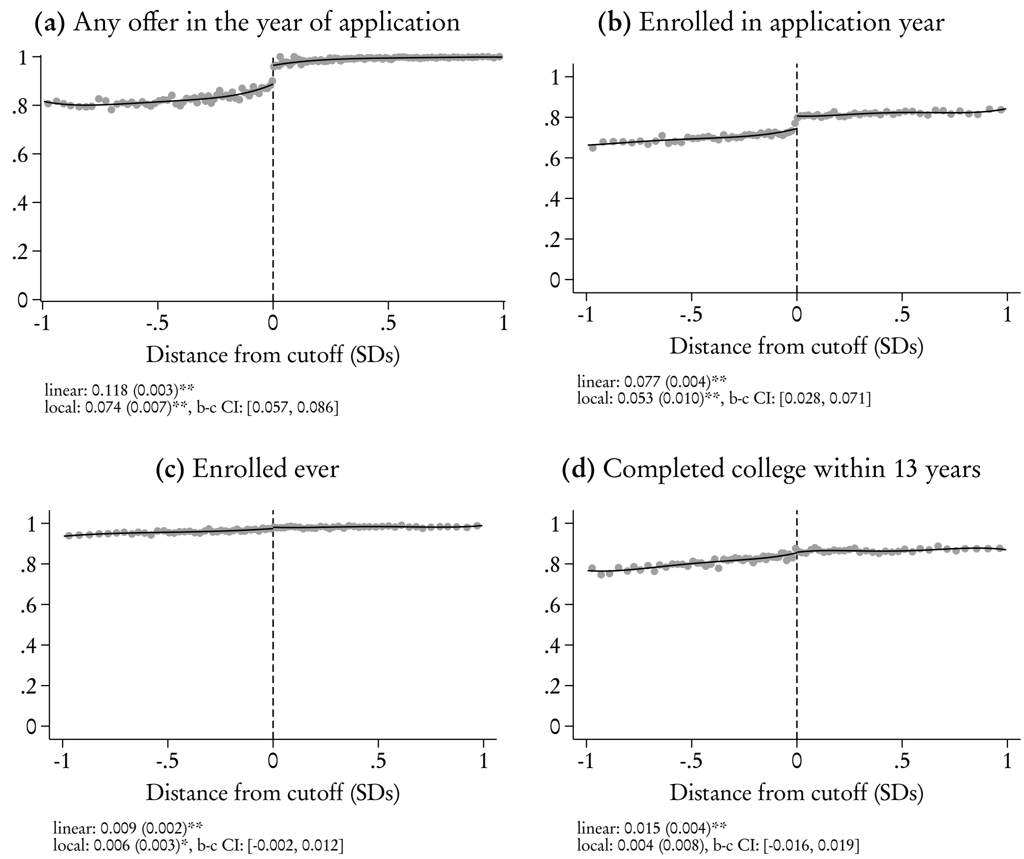
Figure 2. Threshold crossing and sorting

and when they submit their application. Second, the admission cutoffs vary considerably over time, in part because of changes in demand, but also because changes in funding cause variation in the supply of slots. Third, there is limited scope for sorting around the cutoff during the last semester of high school when students do their final exams and apply for post-secondary education. In our setting, the application scores depend on the academic results over all three years in high school, unlike countries in which admission is based only on how well the students do in final year exams or college entrance tests.

Regression-discontinuity estimates on college enrollment and completion

Since students do not appear to sort around the admission thresholds, we can infer the causal effects of threshold crossing by examining how the outcomes change around the thresholds. As above, Figure 3 pools all the programs and admission cutoffs. For each outcome, we plot the unrestricted means in bins and include regression lines on each side of the admission cutoff.

The first two panels of Figure 3 show that having an application score above the threshold increases the likelihood of both receiving an offer and enrolling



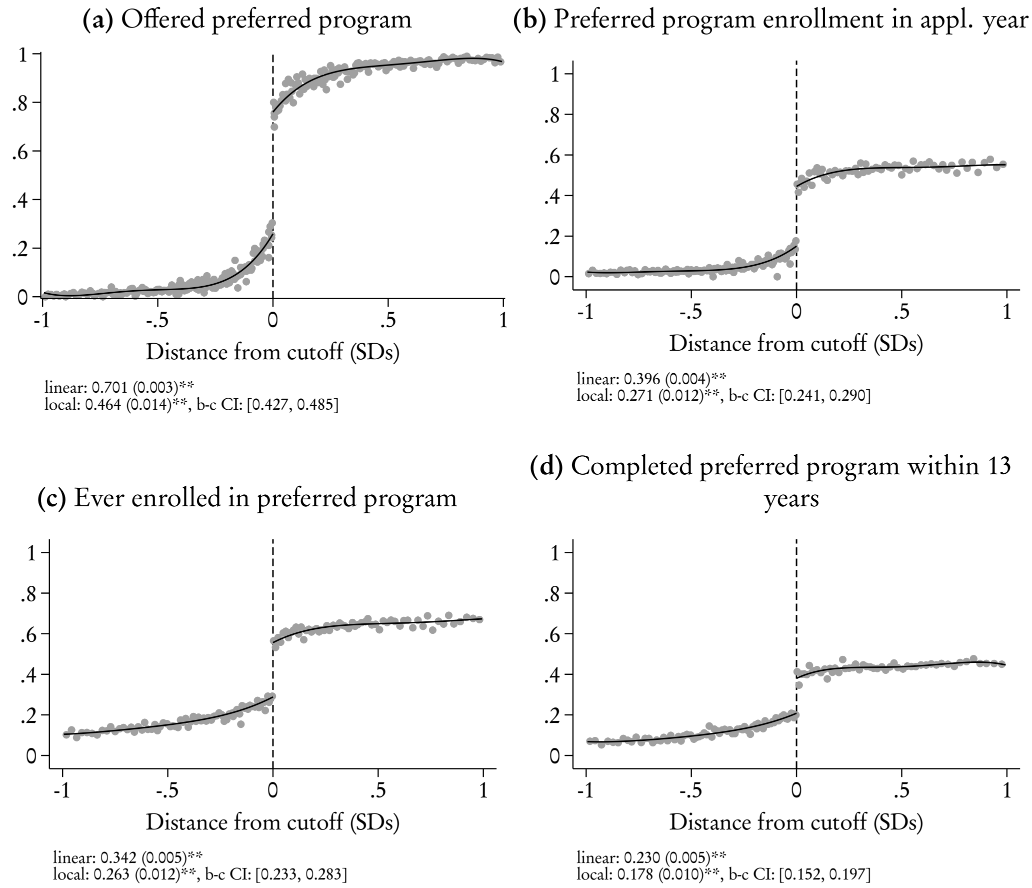
Note: Panel (a) shows share receiving any offer in the application year. Panel (b) shows share enrolling any program in the application year. Panel (c) shows share enrolling any program within 13 years. Panel (d) shows share who have completed any program within 13 years. See notes to Figure 2 for further explanations.

Figure 3. Threshold crossing, college enrollment and completion

in the year of application. The next two panels show that threshold crossing to the preferred program has negligible effects on being ever enrolled in college or completing college. Taken together, these results suggest that threshold crossing only has a small impact on the timing of college enrollment and, more importantly, it does not influence the chances of ever enrolling or ultimately completing college.

Regression-discontinuity estimates on educational choices

While threshold crossing has ultimately little if any impact on enrollment in higher education or college completion, Figure 4 reveals that threshold crossing to preferred program is key for the type of college education that individuals



Note: Panel (a) shows share offered the preferred program in the application year. Panel (b) shows share enrolling in the preferred program in the application year. Panel (c) shows share enrolling in the preferred program within 13 years. Panel (d) shows share who have completed the preferred program within 13 years. See notes to Figure 2 for further explanations.

Figure 4. Threshold crossing, type of college education

choose and complete. Also in this figure, we pool all the programs and admission cutoffs, plot the unrestricted means of the outcomes in bins, and include regression lines on each side of the admission cutoff.

First note that (both here and in Figure 3) there are quite a number of never-takers for enrollment. While we know from the statistics from NUCAS that about 62 percent of those offered a place in a program do accept and show up at the start of the semester, our numbers are somewhat lower because the enrollment data from Statistics Norway are based on enrollment 1–2 months into the academic year by which time some applicants will have already dropped out. Although ultimate enrollment may seem low, it is probably best understood by the fact that the centralized on-line system makes the cost of applying very small. This (and the fact that there are no tuition fees and only small nominal registration cost) is likely to attract a non-negligible number of never-takers.

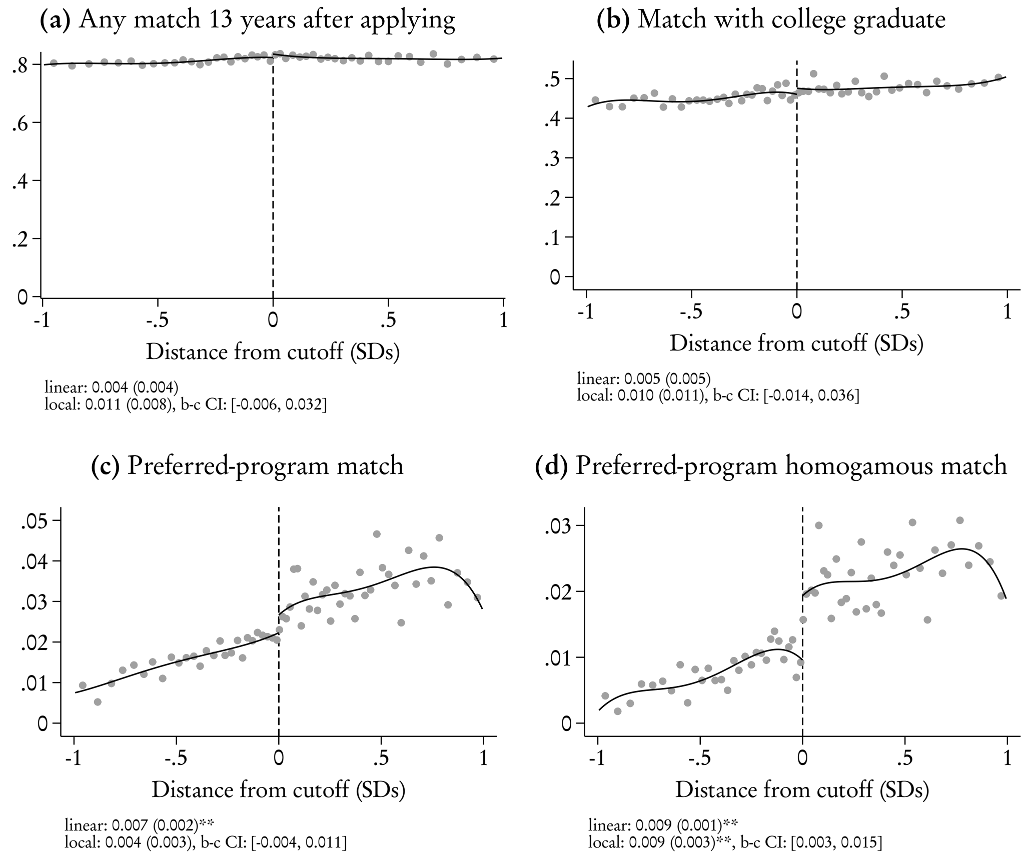
Second, to understand the fuzziness of the RD, note that some applicants that score above the second-round cutoff but below the first-round cutoff do not choose to be on the waiting list (actively, or passively by not responding) and will therefore not get an offer. Also, if applicants decline the final central second-round offers then the institutions may admit more applicants through additional local rounds. Finally, institutions have some limited discretion when evaluating applicants "where there are particularly compelling reasons", e.g. illness or disability, "that make it likely that exam scores do not reflect the applicant's qualifications". This requires applicants to submit appropriate documentation such as for example a medical certificate. The text explaining the process to the applicants makes clear that if this leads to adjustments then they are likely to be only minor.

Panel (a) of Figure 4 shows that crossing the threshold increases the probability of getting an offer for the preferred program by about 46 percentage points. Preferred program enrollment increases initially by 27 percentage points (panel b), and ultimately by 26 percentage points (panel c). By comparison, preferred program completion increases by 18 percentage points (panel d). Thus, we conclude that threshold crossing impacts the type of higher education that people enroll in and complete, not whether people enroll and ultimately complete any college education.

Regression-discontinuity estimates on matching and homogamy

The figures above show that individuals on both sides of the threshold are similar in the rates of college enrollment and completion, but differ in the program in which they get an offer, enroll, and ultimately graduate. Figure 5 examines whether the abrupt changes in program of study at the cutoffs are associated with discontinuous changes in whether and whom one marries where we focus on the first observed match since applying. As above, we pool all the programs and admission cutoffs, plot the unrestricted means of the outcomes in bins, and include regression lines on each side of the admission cutoff.

Panel (a) of Figure 5 shows that crossing the threshold does not change the probability of marriage. Nor does it, as shown in panel (b), change the probabil-

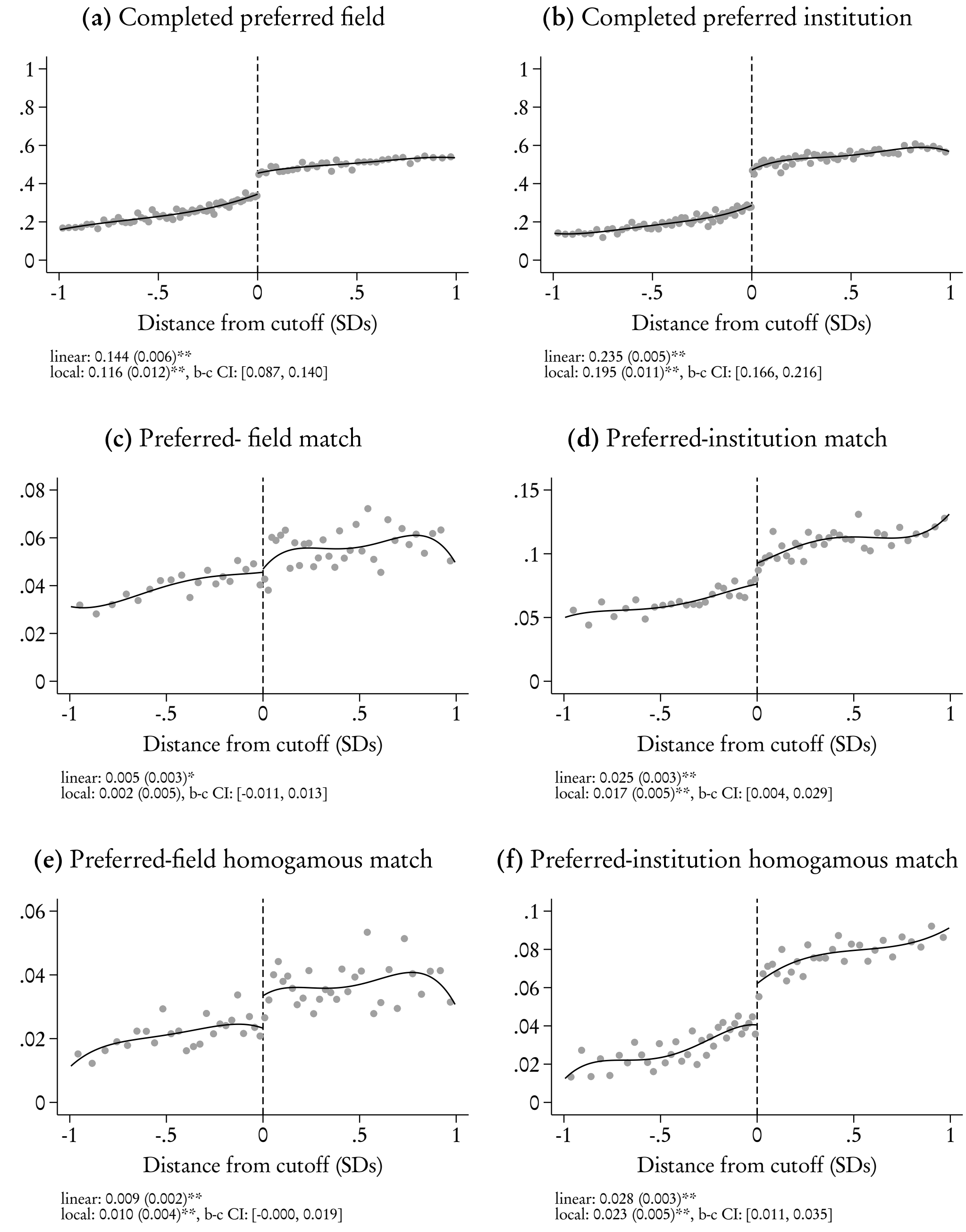


Note: Panel (a) shows share with any match 13 years after applying. Panel (b) shows share with partner with college degree. Panel (c) shows share with partner having completed the applicant's preferred program. Panel (d) shows share who have completed the preferred program and who have a preferred-program match. See notes to Figure 2 for further explanations.

Figure 5. Threshold crossing, marriage and homogamy

ity of matching a college graduate. As evident from panel (c), however, threshold crossing to preferred program makes it more likely to match with someone who has enrolled in and completed exactly the same type of education. At the threshold, this probability increases by 30 percent, compared to the base of 2 percentage points. Likewise, crossing the threshold doubles the chances of a homogamous match where both spouses have eventually completed a degree in the preferred program (panel d).

The regression-discontinuity estimates show that threshold crossing to preferred program changes the type of college education that people enroll in and complete, but it does not affect the probability of enrolling in or completing any type of college. Likewise, threshold crossing to preferred program impacts



Note: See notes to Figure 2, 4 and 5 for further explanations.

Figure 6. Threshold crossing, field/institution of study and homogamy

the propensity to match with someone with a degree in this program. However, it does not affect whether one marries or whether one marries a college graduate. Figure 6 examines these findings in greater detail, breaking down the results for preferred program (defined as the combination of field and institution) into field versus institution of study. We find that threshold crossing has an especially large effect on homogamy with respect to institution.

6 Instrumental variables model

The discontinuities that arise from the college admission process allow us to identify the so-called intention-to-treat effects of crossing the admission thresholds. To estimate the effect of enrollment we now turn to our instrumental variable estimation, which uses the admission thresholds as instruments.

Regression model and parameters of interest

Our IV model uses threshold crossing as an instrument for enrollment in the applicants' preferred program. We will estimate the following first-stage

$$d_i = \pi z_i + x_i'\gamma + u_i \tag{4}$$

where the dependent variable d_i equals 1 if the applicant ever enrolled in the preferred program, field or institution (and zero otherwise). The instrument z_i is the predicted offer for the preferred option, and z_i is therefore equal to one if is the individual's application score exceeds the admission cutoff (and zero otherwise).

The corresponding second stage is as follows

$$w_{it} = \delta_t d_i + x_i'\beta_t + e_{it} \tag{5}$$

where w_{it} is the outcome of interest of individual i in year since applying t. The target of our estimation is the average of δ_t among the compliers who enroll in their preferred program because their application score are just above the admission cutoff to this program and who would not have enrolled otherwise. We use 2SLS with first and second stage equations given by (4) and (5) to estimate δ_t for every year $t = 1, 2, \ldots, 13$. We also decompose the estimated δ_t into the complier average potential outcomes with and without the preferred program. Here we follow Abadie (2003) who shows that with a binary treatment and instrument d and z, and a scalar outcome w one can recover the compliers' mean potential outcome with treatment w^1 from a 2SLS regression of $d \cdot w$ on d instrumented with z. Similarly the compliers' mean potential outcome without treatment w^0 can be recovered from a 2SLS regression of $(1-d) \cdot w$ on $(1-d)$ instrumented with z. This decomposition helps in interpreting the magnitude of the estimated effects.

Identifying assumptions

To identify δ_t, we make three assumptions. The first is that applicants are not able to perfectly sort themselves around the cutoff in order to receive an offer for their preferred field of study. As showed in the previous section, the data supports this assumption. The second assumption is that crossing the admission threshold makes it weakly more likely that an applicant enrolls in the preferred degree. This monotonicity assumption seems plausible in our setting.

The last assumption of our IV model is that threshold crossing affects the outcomes of interest only through the treatment variable. To evaluate this assumption, it is useful to observe that we specify d_i as an indicator for whether an applicant enrolls in the preferred program. This specification alleviates concerns about the exclusion restriction that could arise of one instead used having completed preferred program as the treatment variable. If one uses completion as the treatment variable, the concern would be that individuals who enroll in a given program may be more likely to meet and match with other people in that program, even if they drop out and do not complete the program.

Estimation, empirical specification, and inference

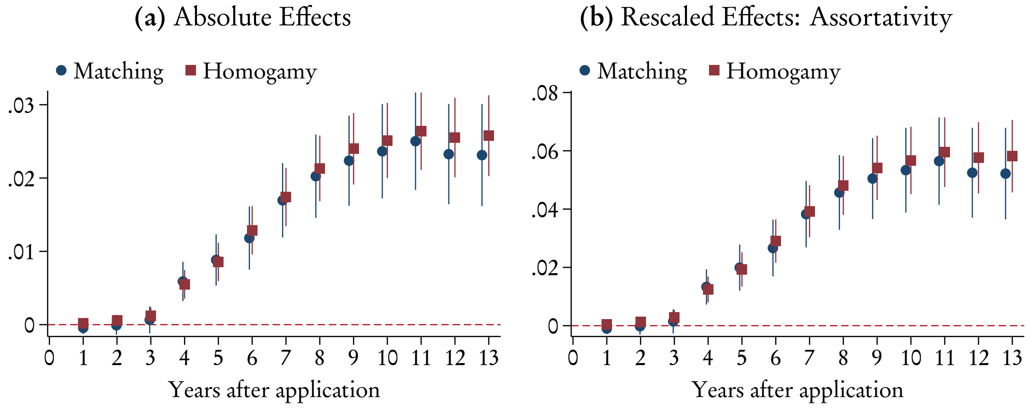
Our estimation approach exploits the fuzzy regression discontinuity design implicit in the admission process described above, where individuals with application scores above the cutoff are more likely to receive an offer from the preferred program. Although the identification in this setup is ultimately local, we use 2SLS because our sample sizes are too small to perform local non-parametric estimation. We need to control for the running variable in x_i to ensure the exogeneity of our instrument. Motivated by the graphical evidence, our baseline specification includes a linear spline in the applicant's application score distance to the admission threshold, thus allowing the slope to vary on each side of the cutoff. To reduce residual variance we also add a set of (pre-determined) controls for gender, application year, and preferred program. About 15–20 percent of the applicants are observed at two margins. To address the dependency this creates between observations we cluster the standard errors at the applicant level. In Figure 11 we present results from several specification checks, all of which support our main findings. We discuss these in more detail in the next section.

Recovering assortativity

As discussed in Section 3, when interpreting measures of educational homogamy it is important to observe that the values they can take on are constrained by the marginal distributions of education among men and women. This motivated rescaling homogamy rates by maximum minus random homogamy $h^m - h^r$. In some comparisons below we also rescale our 2SLS estimates of the effect of enrollment on homogamy. Given that our IV estimates are local to the compliers, we will rescale the estimates by the average scaling factor $h^m - h^r$ for treated compliers (i.e. complier w^1). It is useful to note here that assigning another individual to their preferred option will have a negligible effect of the marginal distribution of potential partners, and therefore have no noticeable effect on maximum possible homogamy and homogamy under random matching. We therefore take h^m and h^r as fixed measures above and below the cutoff. To compute the scaling factor we first construct h^m and h^r based on applicants' preferred education (see the discussion in Section 2.3), and then estimate the average for compliers by estimating 2SLS regressions as in (5), but with $d_i \cdot (h^m - h^r)$ as a dependent variable.

7 College education, homogamy and assortativity

We now present the IV estimates of how the choice of college education affects homogamy and assortativity. We focus on the second-stage estimates. The first-stage estimates are reported in the notes to the figures. F-statistics are never below 2000, suggesting that weak instruments are not a cause of concern.



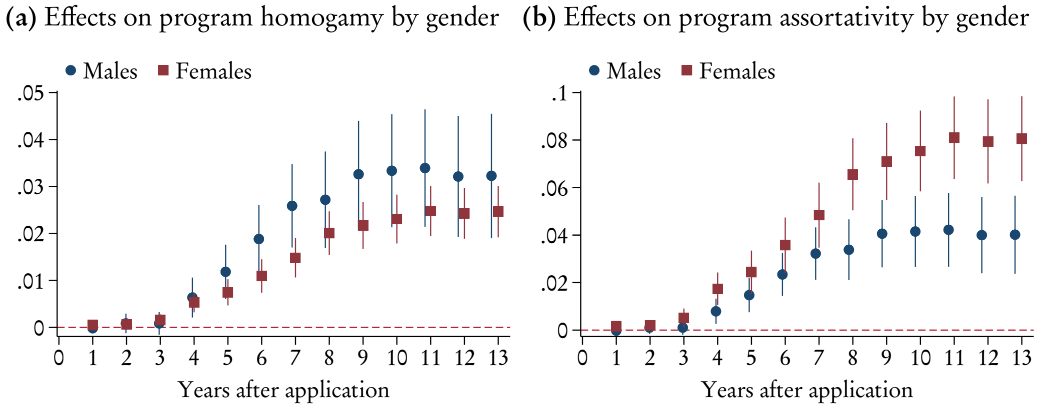
Note: Figure (a) shows 2SLS effects of enrolling in the preferred program, cf. (5), on i) Matching a preferred-program partner and ii) preferred-program Homogamy, i.e. completed preferred program and preferred-program partner. Each estimate comes from a separate regression among applicants at the program margin. Figure (b) shows corresponding effects on effects on preferred-program assortativity, i.e. the effects on preferred-program matching/homogamy rescaled by mean assortativity for treated compliers. The first-stage coefficient is 0.38 (0.004). Error bars indicate 95% CI (standard errors clustered at the applicant level).

Figure 7. IV estimates of the effect of enrollment on program homogamy and assortativity

Effects on homogamy and assortativity with respect to program

Figure 7 reports 2SLS estimates of equations (4) and (5) of the effects of enrolling in the preferred program. We consider two outcome variables. The first is an indicator variable for the applicant matching with someone who holds a degree in the applicant's preferred program. The second is an indicator variable for homogamy with respect to program, that is: both the applicant and the spouse have a degree in the applicant's preferred program. The effects on these two outcome variables can differ only because some applicants may first enroll and then drop out of the preferred program. When interpreting how the estimates change over time, it is useful to observe that the education of the spouse is recorded 13 years after application. As a result, any variation in matching effects over time is due to changes in matching, not changes in the spouse's education.

The results in panel (a) show that enrolling in the preferred program causes a large increase in the probability of matching with someone with a degree in this program. The effects increase over time, plateauing at nearly 2.5 percentage points by year ten. Similarly, enrolling in the preferred program increases the program homogamy rate by about 2.5 percentage points. This is a stark change



Note: Figure (a) shows 2SLS effects of enrolling in the preferred program, cf. (5), on preferred-program homogamy by applicant's sex. Sample sizes range from 50,685 to 49,559 men and 83,918 to 82,647 women. Figure (b) shows corresponding effects on effects on preferred-program assortativity, i.e. the effects on preferred-program matching/homogamy rescaled by mean assortativity for treated compliers. The first-stage coefficient is 0.31 (0.006) for males and 0.41 (0.005) for females . For more details, see note to Figure 7.

Figure 8. IV estimates of the effect of enrollment on program homogamy and assortativity by gender

in educational homogamy given the average program homogamy rate of 0.04. Since the estimated effects on homogamy are nearly identical to those for matching with someone with a degree in the program, we focus, in the remainder of the paper, on the homogamy estimates, which are easier to interpret.

While Panel (a) reports effects on homogamy, panel (b) focuses on the impacts on assortativity. As explained in Section 6, the latter scales the effects relative to the average maximum feasible homogamy for the compliers. As evident from panel (b), the effect on assortativity is considerably larger than the effect on homogamy. By year ten, we find that enrolling in the preferred program increases assortativity by about 6 percentage points. This finding highlights the importance of distinguishing between homogamy and assortativity. The latter is arguably a preferred measure of the degree of sorting, as it takes into account that the maximal attainable homogamy is constrained by the marginal education distributions of men and women.

In Figure 8, we estimate the 2SLS model separately by gender. The point estimates for homogamy are larger for men than women, although the difference is not statistically significant. However, women are overrepresented relative to men in higher education, especially in programs like teaching and nursing, which constrains the maximal attainable homogamy. After taking this into account through the measure of assortativity the effects on women become significantly larger for women as compared to men. For example, the point estimates at 13 years shows that enrolling in the preferred program increases assortativity among women by as much as 8 percentage points compared to 4 percentage points for men.

Effects on homogamy and assortativity with respect to field and institution

Until now we considered homogamy with respect to the applicant's preferred program. A program is defined as a specific field of study at a given institution. We can also consider homogamy with respect to fields or institutions. This is done in Figure 9 which shows separate effects on homogamy and assortativity by preferred program, field and institution.

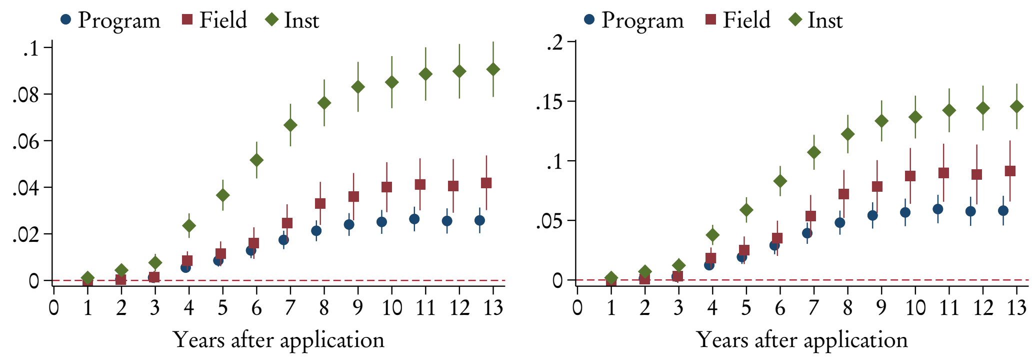
We find that enrolling in the preferred institution has large effects on the probability of homogamy with respect to that same institution. By year ten, enrolling in the preferred institution increases the homogamy with respect to that institution by 9 percentage points. This is a substantial effect on educational homogamy given the average institution homogamy rate of 0.12. By comparison, the effects on homogamy with respect to field are smaller. By year ten, enrolling in the preferred field has increased the homogamy with respect to that field by about 4 percentage points. When interpreting these estimates, however, it is important to observe that the educational distribution of men and women differ less across institutions than across fields and programs. To adjust for this, we also present estimates on the degree of assortativity. These effects are uniformly larger than the effects on homogamy, ranging from about 7 percentage points for programs to 15 percentage points for institutions. This finding shows that the homogamy effects with respect to institution remain larger than those for program and field, even if the effect sizes are measured relative to the maximal attainable homogamy.

Heterogeneity across fields and institutions

Above we documented that the choices of post-secondary education are empirically important in explaining whom but not whether one marries while addressing concerns about self-selection into education based on unobservables.

(a) Effect on program, field and inst. homogamy

(b) Effect on program, field and inst. assortativity



Note: Figure (a) shows 2SLS effects of enrolling in the preferred program, field or institution on preferred-program, field or institution homogamy. Samples are applicants at the margins corresponding to the different measures of homogamy. Sample sizes range from 132,206 to 134,603 applicants at the program margin, from 113,148 to 115,119 at the field margin and from 111,757 to 113,810 at the institution margin. Figure (b) shows corresponding effects on assortativity. The first-stage coefficient is 0.38 (0.004) at the program margin, 0.27 (0.004) at the field margin, and 0.34 (0.004) at the institution margin. For more details, see note to Figure 7.

Figure 9. IV estimates of the effect of enrollment on homogamy and assortativity by program, field of study and institution

Indeed, the magnitude of these estimates are sufficiently large to explain a majority of the strong educational homogamy and assortativity that we observe among the college educated in our data. A natural question to ask is whether these effects of enrollment on homogamy and assortativity are concentrated in certain fields and institutions, or if they apply more broadly.

We investigate this question by estimating the following second-stage equation

$$y_i = \sum_k \delta_k d_{ki} + x_i'\beta + e_i \tag{6}$$

where y_i denotes preferred field/institution homogamy (measured 13 years after applying), and the indicator variables d_{ki} are equal one if applicant i ever enrolled in field/institution k. The corresponding first-stages are as follows

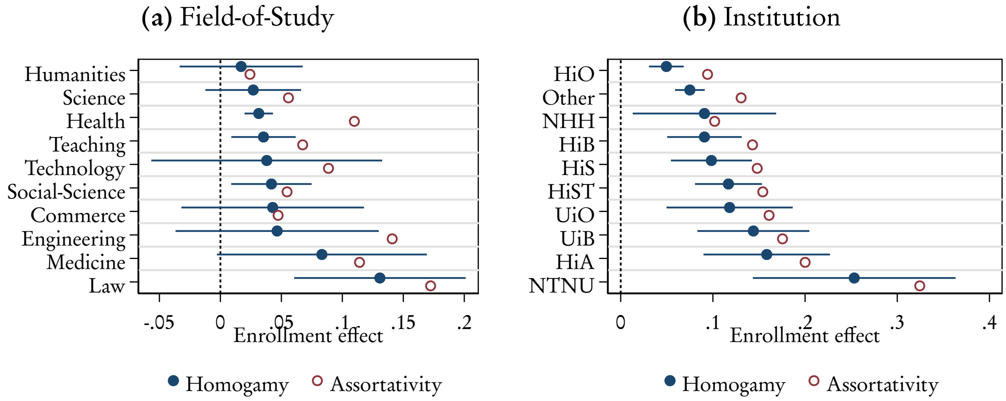
$$d_{ki} = \sum_l \pi_{kl} z_{li} + x_i'\gamma_k + u_{ki} \tag{7}$$

where the instruments z_{li} are equal to one if individual i applied to a program in field/institution k and also had an application score that exceeded the admission cutoff in that program. The control variables in x_i consist of indicator variables for application year, sex and preferred program as well as a linear spline in distance to the application threshold which is allowed to vary across field/institution. Programs are here classified into ten broad fields following Kirkeboen et al. (2016), and we consider the nine largest institutions and a pooled tenth group of the remaining smaller ones. Figure 10 summarizes the findings while the detailed estimation results are reported in Appendix Tables A1 and A2.

The left panel of Figure 10 shows that estimates of enrollment effects on preferred-field homogamy range from about 2 percentage points in the Humanities to about 13 percentage points in Law. While the effects are fairly similar across many fields of study, two fields appear to stand out and those are Medicine and Law where enrollment increases homogamy by 8 and 13 percentage points respectively. The gender composition of students differs however notably across fields. For example, Engineering, Science, and Technology are fields where the large majority of the students is male, while Health and Teaching are female dominated fields. To take this into account, Figure 10 also reports effects of enrollment on assortativity. While we see that accounting for differences in gender composition does matter for fields such as Health and Engineering, it does not attenuate differences across fields.

The right panel in Figure 10 reports estimates broken down at the institution level. The estimates suggest that the enrollment effects on preferred-field homogamy vary substantially across institutions. Effects on assortativity are uniformly larger as the majority of students is female, but taking the gender composition into account does not change the relative order of the effects.

Our results suggest that while differences in gender composition matter for matching, it cannot explain why we observe different effects of enrollment on homogamy across both fields and institutions. Are the enrollments effects larger for more selective fields and institutions? And are they larger when earnings are higher? In Appendix Table A1 we also report average GPA – a measure of selectivity – and earnings across the different fields and institutions. Average earnings differ by nearly a factor two when going from the lowest ranked to highest ranked field and institution. We also observe nearly a full standard deviation difference in average GPA across both fields and institutions. When we relate



Note: These figures show 2SLS effects of enrolling in a given preferred-field (Figure a) or institution (Figure b) on preferred-field or institution homogamy or assortativity. Samples are applicants at the margins corresponding to the different measures of homogamy. Sample size is 111,397 at the field margin and 110,507 at the institution margin. The first-stage coefficients are reported in Appendix Tables A1 and A2. For more details, see note to Figure 7.

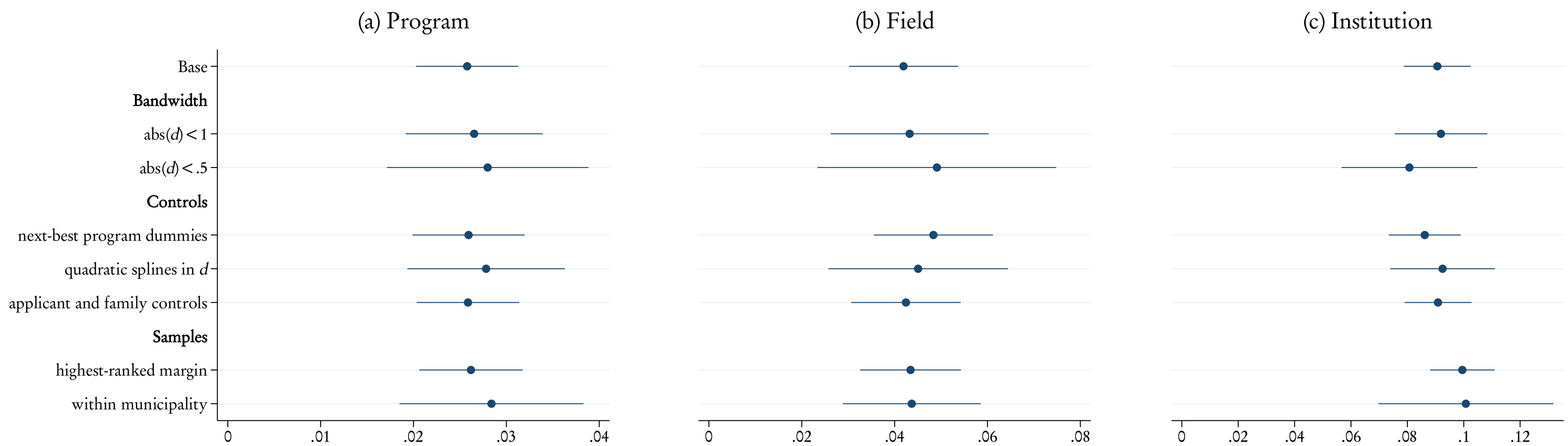
Figure 10. Effect of enrollment on homogamy and assortativity – By preferred field-of-study and institution

these measures with the enrollment effects we uncover correlations in the neighborhood of 0.5 both at the institution and field margin, and these correlations are unchanged when we account for differences in gender composition. Thus, we may conclude that the effects of enrollment on educational homogamy and assortativity vary systematically across fields and institutions, and tend to be larger in more selective and higher paying fields and institutions.

Specification checks

In Figure 11, we present results from a battery of specification checks. For brevity, we only report second stage estimates for outcomes measured in the last year of our data, 13 years after applying.

In each graph of Figure 11, the top line reports our baseline estimate, corresponding to those presented in Figure 9. In the first two specification checks, we show that our estimates are robust to choice of bandwidth and do not materially change if we change the bandwidth in the regression discontinuity estimation. In the next three specification checks we show that our estimates are robust to adding more controls, including indicator variables for next-best program, quadratic splines in the distance to the cut-off, and additional pre-determined



Note: Each figure shows 2SLS estimates of the effects of ever enrolling in the preferred education on preferred-education homogamy, i.e. completed preferred education and preferred-education partner, cf. (5). Panel (a) reports effects on preferred-program homogamy for applicants at a margin between two different programs, (b) shows effects on preferred-field homogamy for applicants at field margins and (c) shows effects on preferred-institution homogamy for applicants at institution margins. Homogamy is observed 13 years after application. Each estimate is from a separate regression with a given sample and set of controls, as indicated in the figure. Error bars indicate 95% CI (standard errors clustered at the applicant level).

Figure 11. Robustness of enrollment effects on homogamy 13 years after enrolling

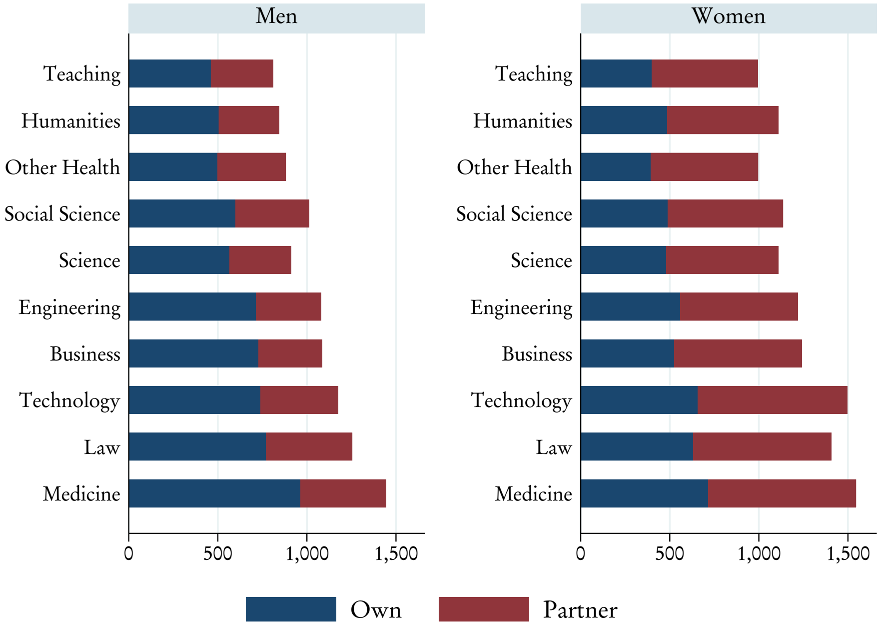
variables for applicant and family background (including dummies for age at application, municipality of residence at age 16 and dummies for whether parents are immigrants and have higher education). As some applicants in our sample are observed at observed at two margins (see the discussion in Section 2), we include a specification check where we only keep the most-preferred margin of each applicant. Reassuringly, the estimates do not materially change. One issue, which is particularly relevant for the interpretation of the institution effects, is whether the treatment is confounded by geography. To investigate this question the last rows in Figure 11 restrict the sample to applicants who are on the margin between program, fields or institutions within the same municipality. This ensures that the reported effects only capture the consequences of within municipality assignments and the results show that we can rule out that (changes in) locality explain our results.

The importance of homogamy for family earnings

The results above showed that while the choice of post-secondary education does not matter for whether one marries, it is empirically important in explaining whom one marries. We now discuss the potential earnings implications of these results.

We begin this investigation with an analysis of how individual and family earnings vary with choice of post-secondary education. To do so, we estimate complier average potential outcomes across different fields of study following the 2SLS estimation outlined in Section 6 for own earnings y_i^{own} and partner earnings y_i^p. We perform this analysis separately for males and females. We focus on homogeneity by field of study because in Norway the effect on earnings from attending a more selective institution tends to be relatively small compared to payoffs to field of study as shown in Kirkeboen et al. (2016).

Figure 12 starts out by reporting average potential earnings and partner earnings for compliers. We find that the average potential family earnings of the compliers who enroll in medicine, law and technology are much higher than the average potential family earnings of those who enroll in teaching and humanities. This is true both for men and women. For women, however, a majority of the potential family earnings can be attributed to their partner's potential earnings. Partners' potential earnings exceed own potential earnings even for women enrolled in the highest paying fields, such as medicine, law, and tech-

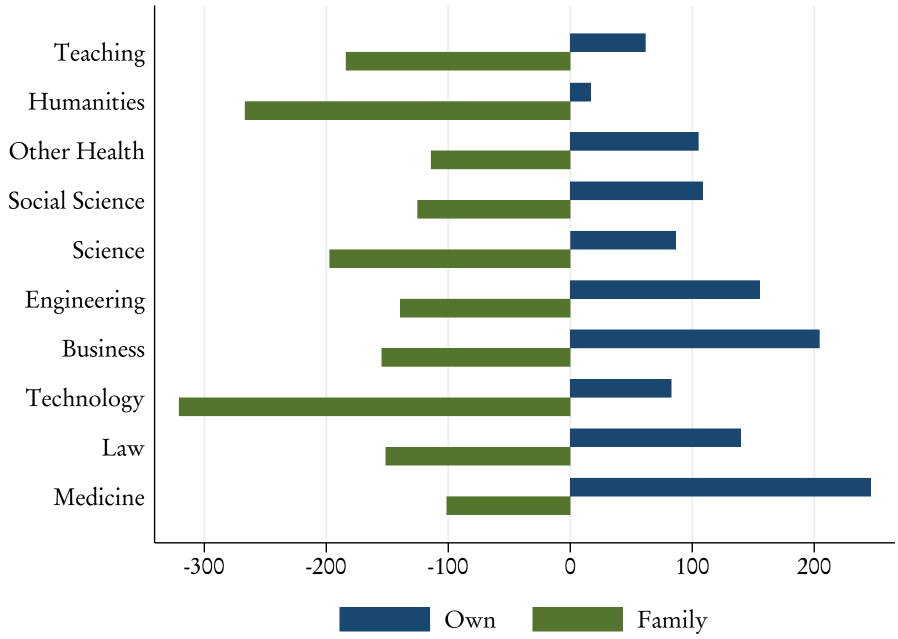


Note: Each figure reports 2SLS estimates of average own and partner potential yearly earnings for compliers cf. (5). For own and partner's earnings and for each sex we estimate a regression model instrumenting enrollment · preferred field with above cut-off · preferred field. The regressions control for a spline in the distance to the cut-off, preferred field and application year. The sample is all matched applicants on a margin between different fields (the next-best alternative may be no college) for whom own and partner's earnings are observed (missing earnings data excludes 1.3% of matched applicants). Earnings are observed 13 years after application and measured in 1000 NOK, corresponding to about 114 USD.

Figure 12. Family potential earnings (own + partner's potential earnings), by sex and preferred field

nology. By comparison, partners' earnings make up a relatively small share of family earnings for men who enroll in these fields. The gender gap in earnings by field of study reverses if one looks at family earnings instead of individual earnings. This is shown in Figure 13. We see that in every field of study, men have a higher individual potential earnings but lower family potential earnings than women.

These findings point to the potential importance of the choice of post-secondary education for family earnings. This does not necessarily imply, however, that the enrollment effects on homogamy have important effects on family earnings. As a first step to investigate the importance of homogamy, we decompose av-



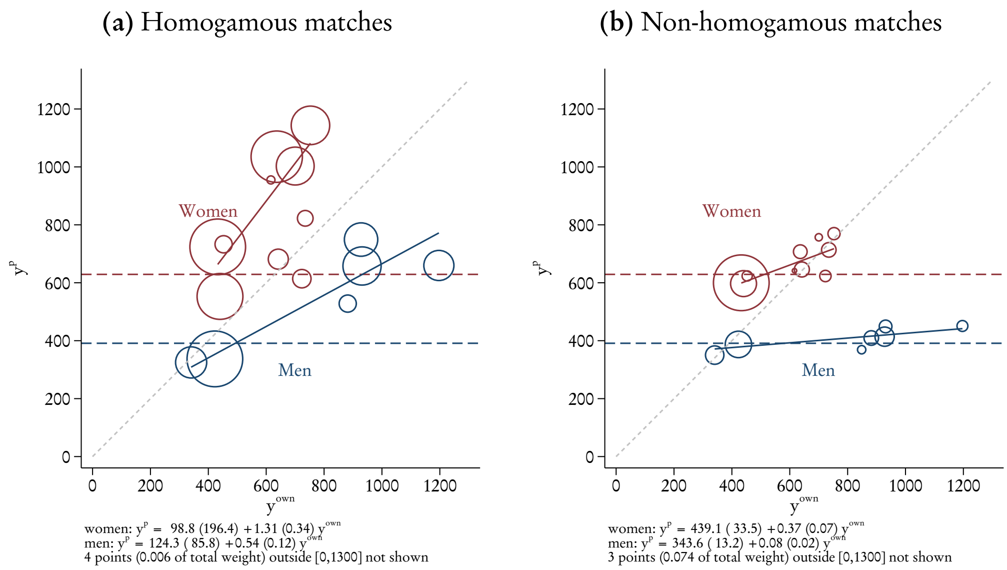
Note: Family potential earnings is own + partner's potential earnings, see the note to Figure 12 for details on earnings measure, estimation and sample.

Figure 13. Gender difference (male - female) in own and family potential earnings, by field

erage potential partner earnings into average potential partner earnings for homogeneously and non-homogeneously matched compliers departing from the following identity

$$y_i^p = y_i^p h_i + y_i^p (1 - h_i) \tag{8}$$

where h_i is a preferred-field homogeneous marriage indicator. We estimate average potential partner earnings for homogeneously matched compliers by estimating equation (5) with $y_i^p h_i$ as the dependent variable, and average potential partner earnings for non-homogeneously matched compliers using $y_i^p(1-h_i)$ as the dependent variable. Given (8), these estimates will sum to the average potential partner incomes for the compliers reported in Figures 12 and 13. Having estimated potential complier averages $\widehat{y^p h}$ and $\hat{h}$ for each preferred field, we recover average partner's earnings of compliers in homogamous matches as $\hat{y}^{p,h} = \widehat{y^p h}/\hat{h}$. As above we perform these estimations separately for men and women, and then repeat the same analysis for own earnings y_i^{own}. Figure 14 reports the results.



Note: Figures show partner's potential earnings vs. applicants potential earnings by preferred field, sex and whether match is homogamous. Potential earnings in homogamous matches are calculated as total potential earnings in homogamous matches divided by share homogamous matches, and similarly for non-homogamous matches. Horizontal line is average partner potential earnings. Solid lines are linear fits (using all fields) and weighted with # compliers · share homogamy for treated compliers in homogamous matches and with # compliers · (1 - share homogamy) for treated compliers in non-homogamous matches.

Figure 14. Potential partner's earnings (y^p) vs. own potential earnings (y^{own}) for homogamously and non-homogamously matched compliers

Figure 14a contrasts, for every field-of-study, average potential partner earnings against own potential earnings for homogeneously matched compliers. The estimates for women are marked by the red circles, and the estimates for men are blue where the size of the symbol is proportional to the number of compliers in homogamous matches. We observe that the estimates for women lie above the 45-degree line which indicates that partner potential earnings is on average higher than own potential earnings. We see the opposite for men. The solid estimated regression lines show the strong positive association between potential partner and own earnings across fields for both women and men. The estimated slope is 1.3 for women and 0.5 for men. The typical benchmark to quantify the importance of homogamy is to compare to random sorting, which is indicated by the horizontal dashed lines. This comparison suggests that homogamy is important for family income.

The counterfactual of unconditional random sorting arguably puts an upper bound on the importance of homogamy. A weaker counterfactual assumption is that in absence of homogamy the compliers to a given field of study would have matched like compliers to the same preferred field who find themselves in non-homogamous marriages. Figure 14b relates potential partner earnings to own earnings for these non-homogeneously matched compliers. We observe again that women have higher- and men lower-earning partners, but the relationship between own and partner earnings is now much attenuated for both sexes. The estimated slope is 0.37 for women and thus much closer to the benchmark of random sorting than in Figure 14a. For non-homogeneously matched male compliers the relationship is even weaker, at 0.08.[10]

Taken together, the results in Figure 14 suggest that the effects of enrollment on educational homogamy are economically important as the potential earnings of partners in homogamous matches are not only materially different from the benchmark of random sorting, but also from the potential earnings of partners in non-homogamous matches. In other words, non-homogamous partners are far from perfect substitutes to homogamous partners in terms of potential earnings.

8 What determines homogamy among the college educated?

Above, we documented that the choices of college education are important determinant of educational homogamy and assortativity. Motivated by this finding, we now examine the ways in which this educational homogamy and assortativity arise.

Our point of departure is the textbook models of marriage markets where educational homogamy and assortativity arise as equilibrium outcomes of preferences for spousal education given the marginal distributions of education of men and women (see e.g. Choo and Siow, 2006; Chiappori et al., 2017).[11] However,

[10]To investigate to what extent these findings are driven by labor supply responses we also performed the same analyses on hourly wages (not reported here, and available on request) and find qualitatively very similar results.

[11]The literature discusses two types of preferences. One possibility is that individuals have horizontal preferences for same-education spouses. For example, doctors may have preferences for matching with doctors and lawyers may have preferences for matching with lawyers. The result of such preferences may be that both lawyers and doctors tend to marry spouses with the same type of education. Another possibility is that individuals have vertical preferences across education types, giving a uniform ranking of spouses with ranks monotonically related to the spouse's education. For example, both lawyers and doctors may prefer to marry doctors. Yet,

ever, these models rely on assumptions that may not hold in practice. One issue is that individuals may match on traits correlated with education, and not (only) education per se. These traits may be unobserved to the econometrician, such as innate ability or tastes.[12] Another issue is the assumption of a large and frictionless market. In reality, search frictions are likely to depend on educational choices, as it is easier to meet people with the same education, both while in school and later at work. The goal of our analysis is to address and investigate these issues while considering the question of why college graduates tend to marry spouses with the same type of education.

8.1 *Importance of traits determined prior to enrollment*

The discontinuities that arise from the admission process allow us to eliminate any correlation between educational choices and pre-determined unobserved heterogeneity in preferences, ability, and other factors. This implies that our finding of large impacts on homogamy and assortativity reflect the educational choices that individuals make, and not their pre-determined traits. However, this does not imply that predetermined factors do not matter for the educational homogamy and assortativity we observe in the data. To shed light on the potential importance of such pre-determined traits, we estimate the *potential* homogamy of the compliers to our instrument if they did not enroll in the preferred program (i.e. the complier mean potential outcome without treatment y^0, as described in Section 6).

Table 5 reports random matching rates and the effect of enrollment on different types of homogamy (preferred program, field and institution). For reference the first two columns repeat the average and random homogamy rates from Table 4.The third column reports the counterfactual level of matching y^0 that we would have observed for the compliers if they would not have enrolled in their preferred program option. This counterfactual level of matching reflects the role of ex-ante traits of the individuals. The fact that counterfactual levels of matching are relatively close to random levels of matching indicates that pre-

in equilibrium, homogamy may arise as doctors are more attractive in the marriage market as compared to lawyers.

[12]See Chiappori et al. (2018) for a model of the marriage market where individuals match on human capital, which depends on education but also on innate ability. They allow the latter to be unobserved to the econometrician, and use wage and labor supply dynamics to recover the joint distribution of education and ability and therefore of human capital.

Table 5. Comparison of homogamy rates

	Observed homogamy	Random homogamy	Complier Y^0		Effect of enrollment	
Preferred						
- Program	0.046	0.003	0.011	(0.001)	0.023	(0.004)
- Field	0.062	0.014	0.028	(0.003)	0.027	(0.008)
- Institution	0.121	0.017	0.036	(0.003)	0.080	(0.008)

Note: Observed matching is average program, field or institution homogamy rates in the sample (see Table 4). Random matching is constructed by randomly assigning partners given the population of potential partners for each gender and application year. Complier Y^0 is the estimated level of preferred-education matching for untreated compliers, while effect of enrollment is the 2SLS estimate of the effect of enrolling. Standard errors in parentheses.

determined traits play a modest role in explaining observed homogamy rates. By comparison, the magnitude of the enrollment effects in the final column suggests that educational choices per se can explain nearly all the educational homogamy and assortativity among the college educated (at least among the compliers).

8.2 *What are the key determinants of educational homogamy?*

The results above showed that enrollment can account for a large share of educational assortativity among the college educated. But what are the mechanisms? To help understand the determinants of educational homogamy, we now perform a series of decompositions. These decompositions reveal the following: a) the effect of enrollment on homogamy with respect to field is completely explained by within-institution matches; b) with respect to institution one third of the effect of enrollment on homogamy is explained by matches that are specific to the field of study, while the remainder are across-field matches; c) nearly all the effect of enrollment on homogamy is explained by within cohort matches; d) only a small part of the effect of enrollment on homogamy can be explained by matches that happens within the same workplace. Taken together, these results suggests that colleges are effectively local marriage markets, mattering greatly for whom one marries, not because of the pre-determined traits of the students that are admitted but as a direct result of attending a particular institution at a given time.

Institution and field of study Matches that are homogamous with respect to field (F) are either field-homogamous at the institution level ($F \times I$) or field-homogamous across institutions ($F \times !I$),

$$\Pr(F) = \Pr(F \times I) + \Pr(F \times !I).$$

By estimating the effects of enrollment on each component of this identity, we can decompose the effect on field-homogamous matching into an effect on matching on institution within the same field and an effect on matching on institutions across fields. The left panel of Figure 15 reports the results. We find that the effect of enrollment on homogamy by field is completely explained by within-institution matches.

By the same argument, we can take advantage of the identity that homogamous matches with respect to institution are either institution-homogamous at the field level ($F \times I$) or institution-homogamous across fields ($F \times !I$),

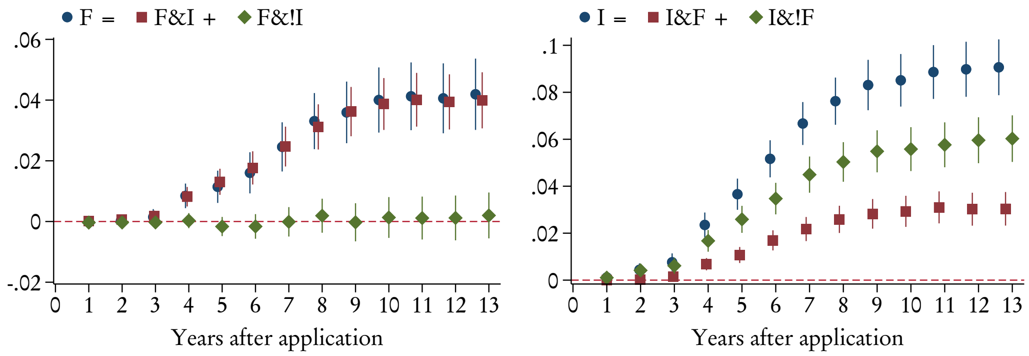
$$\Pr(I) = \Pr(I \times F) + \Pr(I \times !F),$$

and perform a similar decomposition of the effect of education on institution-homogamous matching into the effect matching on fields within the same institution and an effect on matching on fields across institutions. These results are reported in the right panel of Figure 15. We find that about one third of the effect of enrollment on homogamy by institution is explained by within-field matches, while the remainder is due to across-field matches.

Time of enrollment Most programs in Norwegian post-secondary education consist of mandatory and elective courses, where the mandatory courses are usually taken with students from the same cohort. We therefore expect that students are more likely to meet potential partners from their own cohort than from other cohorts. To further develop this notion, we take advantage of the identity that that homogamous matches with respect to institution are either homogamous with respect to enrollment in the same application year (T) or not ($!T$):

$$\Pr(I) = \Pr(I \times T) + \Pr(I \times !T).$$

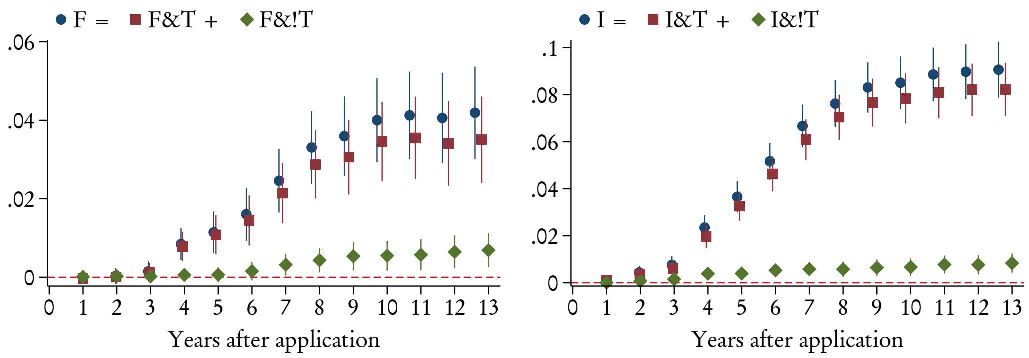
(a) Preferred-field homogamy within and across institutions **(b)** Preferred-institution homogamy within and across fields



Note: Figure (a) decomposes the 2SLS effect of enrollment on preferred-field homogamy (F) into effects on matching field and institution ($F\&I$) and effects on matching field but not matching institution ($F\&!I$). Figure (b) similarly decomposes the effect of enrollment on preferred-institution homogamy (I) into effect on same-field institution homogamy ($I\&F$) and not-same-field institution homogamy ($I\&!F$). Error bars indicate 95% CI (standard errors clustered at the applicant level).

Figure 15. Decomposing the effect of enrollment on homogamy by field of study (F) and institution (I)

(a) Preferred-field homogamy by cohort **(b)** Preferred-institution homogamy by cohort

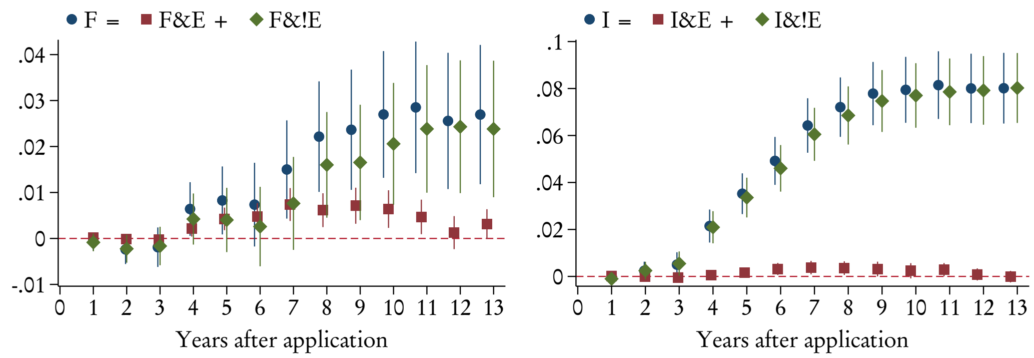


Note: Figure (a) decomposes the 2SLS effect of enrollment on preferred-field homogamy (F) into effects on field-homogamy and cohort (same time of enrollment) ($F\&T$) and effects on field homogamy but different cohort ($F\&!T$). Figure (b) similarly decomposes preferred-institution homogamy (I). Error bars indicate 95% CI (standard errors clustered at the applicant level).

Figure 16. Decomposing the effect of enrollment on homogamy (F, I) by cohort (T)

(a) Preferred-field homogamy within and across employers

(b) Preferred-institution homogamy within and across employers



Note: Figure (a) decomposes the 2SLS effect of enrollment on preferred-field homogamy (F) into effects on field homogamy and overlapping employment history ($F\&E$) and effects on field-homogamy but never same employers at the same time ($F\&!E$). Figure (b) decomposes preferred-institution homogamy (I). Error bars indicate 95% CI (standard errors clustered at the applicant level).

Figure 17. Decomposing the effect of enrollment on homogamy (F, I) by employer-homogamy (E)

By estimating the effects of enrollment on the components of this identity, we can decompose the effect on institution-homogamous matching into an effect on matching on institution within the same cohort and an effect on matching on institutions across cohorts. A similar decomposition follows for field-of-study (F). Figure 16 reports the results. We find that nearly all the effect of enrollment on homogamy is explained by within cohort matches. Not only do colleges act as marriage markets, but search appears also to be local within institutions.

Workplace While educational choices matter for whom you meet in college, they may also affect matching after graduation through work. Thanks to our employer-employee data, we can not only track people through the education system, but also into the labor force and more specifically across firms. We can therefore see whether individuals who match have overlapped at the same employer (before we observe the match and within 13 years since applying). Taking advantage of this data, we can compute the components of the identity that that homogamous matches with respect to institution (or field) are either homogamous with respect to having had the same employer (E) or not ($!E$):

$$\Pr(F) = \Pr(F \times E) + \Pr(F \times !E)$$

By estimating the effects of enrollment on the components of this identity, we can decompose the effect on institution or field homogamous matching into an effect within the same employer and an effect on matching on institutions across employers. Figure 17 reports the results from this decomposition. The results suggest that only a small part of the effect of enrollment on educational homogamy can be attributed to matches within the same workplace.

8.3 Matching functions and market size and tightness

Above, we documented that colleges are effectively local marriage markets and key determinants of educational homogamy among the college educated. The effects on educational choices on homogamy are local to institution, and having overlapped at college. These findings seem at odds with frictionless matching models where it is typically assumed that each person has free access to the pool of all potential partners, with perfect knowledge of the characteristics of each of them—and vice versa. This assumption leads matching models to disregard the cost of acquiring information about potential matches as well as the role of meeting technologies. In contrast, frictions are paramount in search models of the marriage market. In these models, each individual is typically assumed to sequentially meet one person of the opposite gender; after such a meeting, both individuals must decide whether to settle for the current mate or continue searching. To summarize and quantify this complicated search process in terms of a few variables, matching functions have proven useful (see e.g. Petrongolo and Pissarides, 2001). Building on Barnichon and Figura (2015), we specify and estimate market-specific matching functions that helps investigate the impact of local market tightness and size on educational homogamy.

Empirical model We allow the marriage market to be segmented in submarkets, where each submarket k is described by a measure of market size (number of students), and a measure of market tightness (sex ratio). Given that our interest in colleges as marriage markets, we consider three alternative specifications of these submarkets: programs, fields or institutions. In each specification, we empirically model the probability of a preferred-education homogamous match ($y_i = 1$) as depending on enrollment decision and the size and tightness of the market

$$
\begin{aligned}
y_i = \delta d_i + \alpha_1 d_i \log r_i + \alpha_2 d_i \log n_i \\
+ \alpha_3 (1 - d_i) \log r_i + \alpha_4 (1 - d_i) \log n_i + x_i' \phi + u_{it}
\end{aligned} \tag{9}
$$

where d_i equals one if i enrolls in her preferred program (field or institution), $\log r_i$ is the log of the sex-ratio (# own sex / # other sex) in the program (field or institution) applicant i enrolled in, and $\log n_i$ is the log of the size (# other sex) of that program (field or institution). We compute r_i and n_i for each combination of education type (program, field or institution), enrollment year, and gender. The regression model is estimated separately by gender, allowing each coefficient to vary freely across men and women.

When estimating (9), we address concerns about selection and correlated unobservables by using the predicted offer to instrument enrollment d_i. The predicted offer equals one if is the individual's application score exceeds the admission cutoff (and zero otherwise). Depending on whether the individual's score is higher (lower) than the admission cutoff, the individual is also predicted to be exposed to the peers in the preferred program (field or institution). This allows us to instrument the size and sex ratio of the attended program (field or institution) with the size and sex ratio of the predicted program (field or institution), namely the size and sex-ratio of the preferred program if the applicant's application score is above the admission cutoff, and the size and sex-ratio of the next-best program if the applicant's application score is lower. In both the first and the second stages, we control for the running variable by including a linear spline in the applicant's application score distance to the admission threshold, thus allowing the slope to vary on each side of the cutoff. We also add controls for gender, application year, and preferred and next-best program. Given these controls, we are identified from temporal variation in (predicted) size and sex ratios within each program (field or institution).

Empirical results Table 6 reports the IV estimates of equation (9) separately by gender. The tests of the rank condition reported in the table show that weak instruments are of little if any concern. As evident from Table 6, enrollment in

Table 6. IV estimates of effects of enrollment on program/field/institution homogamy by sex-ratio and size of the program/field/institution

	Women			Men		
	Program	Field	Inst	Program	Field	Inst
d_i	0.046	0.065	0.056	0.036	0.080	0.085
	(0.012)	(0.016)	(0.012)	(0.012)	(0.012)	(0.016)
$d_i \cdot \log r_i$	-0.015	-0.039	-0.030	-0.010	-0.025	0.046
	(0.005)	(0.014)	(0.023)	(0.006)	(0.008)	(0.020)
$d_i \cdot \log n_i$	0.003	-0.004	0.020	0.004	-0.020	0.036
	(0.003)	(0.007)	(0.011)	(0.005)	(0.007)	(0.010)
$(1 - d_i) \cdot \log r_i$	0.005	0.023	0.010	0.001	0.012	0.033
	(0.004)	(0.009)	(0.013)	(0.006)	(0.009)	(0.024)
$(1 - d_i) \cdot \log n_i$	-0.003	0.003	-0.007	0.004	0.012	-0.005
	(0.003)	(0.005)	(0.009)	(0.004)	(0.006)	(0.009)
Mean dep. var.	0.016	0.028	0.052	0.026	0.040	0.075
Rank test (χ^2_1)	150.1	32.6	179.5	96.7	18.5	87.5
N	53,568	39,660	41,732	28,901	22,629	22,651

Note: R_i = Sex-ratio = # own sex /# other sex. The reported rank test is the LM version of the Kleibergen and Paap (2006) rk LM-statistic. Clustered standard errors in parentheses.

the preferred program – evaluated at the average program size and sex-ratio in the sample – increases the likelihood of a homogeneous match by 3 percentage points for men and about 4 percentage points for women, which is in line with our estimates in Section 7. Both for men and women we find that the tighter the market (a higher sex-ratio r_i) in the preferred program, the lower the likelihood of a homogamous match. There is, however, no indication that the size (n_i) of the preferred program matters significantly for matching.

The estimates for institution and field homogamy are broadly consistent with those at the program level. Enrolling in the preferred institution or field has a large impact on the likelihood of a homogamy on that margin. While more market tightness also reduces the likelihood of an preferred institution and field homogamous match, there appear to be increasing returns to scale at the institution level. For institution and field homogamy, we also find that market tightness in the alternative, next-best education matters for homogamy. To see this, note that a positive coefficient on $(1 - d_i) \cdot \log r_i$ means that individuals are more likely to marry someone in the preferred education if there are relatively few students of the opposite sex in the next-best education in which they enrolled.

Although the enrollment effects in Table 6 are larger for men than for women, we cannot reject that they are equal. This is also true if we evaluate them at a sex-ratio of 0.5 and the average program size of 200. Furthermore, we cannot reject that the effects of market tightness and market size are also similar across gender.

9 Conclusion

How does the choice of college education affect whether and whom one marries? Why are college graduates so likely to marry someone within their own institution or field of study? We answered these questions using administrative data for Norway's post-secondary education system. A centralized admission process created instruments for choice of college education from discontinuities that effectively randomize applicants near unpredictable admission cutoffs into different institutions and fields of study. Thus, differences in marriage market outcomes across these applicants are due to the institution or field to which they are exogenously assigned, as opposed to their pre-existing abilities, preferences, or family background.

Our instrumental variables estimates are summarized with six broad conclusions. First, the choices of post-secondary education are empirically important in explaining whom but not whether one marries. Indeed, the magnitude of these estimates are sufficiently large to explain a majority of the strong educational homogamy and assortativity that we observe among the college educated in our data. Second, enrolling in a particular institution makes it much more likely to marry someone from that institution. These effects are especially large if individuals overlapped in college, and they are sizable even if the spouse studied a different field. Third, enrolling in a particular field increases the chances of marrying someone within the field insofar the individuals attended the same institution. By contrast, enrolling in a field makes it no more likely to marry someone from other institutions with the same field. Fourth, the effects of enrollment on educational homogamy and assortativity vary systematically across fields and institutions, and tend to larger in more selective and higher paying fields and institutions. Fifth, only a small part of the effect of enrollment on educational homogamy can be attributed to matches within the same workplace.

Lastly, the effects on the probability of marrying someone within their institution and field vary systematically with cohort-to-cohort variation in sex ratios within institutions and fields. This finding is at odds with the assumption in canonical matching models with large and frictionless marriage markets.

References

Abadie, A. (2003). Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113(2):231–263.

Artmann, E., Ketel, N., Oosterbeek, H., and van der Klaauw, B. (2018). Field of study and family outcomes. IZA Discussion Paper No. 11658.

Barnichon, R. and Figura, A. (2015). Labor market heterogeneity and the aggregate matching function. *American Economic Journal: Macroeconomics*, 7(4):222–49.

Belot, M. and Francesconi, M. (2013). Dating preferences and meeting opportunities in mate choice decisions. *Journal of Human Resources*, 48(2):474–508.

Bičáková, A. and Jurajda, Š. (2016). Field-of-study homogamy. *CERGE-EI Working Paper Series*, (561).

Blossfeld, H.-P. (2009). Educational assortative marriage in comparative perspective. *Annual Review of Sociology*, 35:513–530.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, 14(4):909–946.

Chiappori, P.-A. (2020). The theory and empirics of the marriage market. *Annual Review of Economics*, 12:547–578.

Chiappori, P.-A., Dias, M. C., and Meghir, C. (2018). The marriage market, labor supply, and education choice. *Journal of Political Economy*, 126(S1):S26–S72.

Chiappori, P.-A., Salanié, B., and Weiss, Y. (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review*, 107(8):2109–67.

Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of Political Economy*, 114(1):175–201.

Eika, L., Mogstad, M., and Zafar, B. (2019). Educational assortative mating and household income inequality. *Journal of Political Economy*, 127(6):2795–2835.

Han, S. and Qian, Y. (2020). Concentration and dispersion: School-to-work linkages and their impact on occupational assortative mating. *Social Science Journal*.

Hitsch, G. J., Hortaçsu, A., and Ariely, D. (2010). Matching and sorting in online dating. *American Economic Review*, 100(1):130–63.

Juhn, C. and McCue, K. (2017). Specialization then and now: Marriage, children, and the gender earnings gap across cohorts. *Journal of Economic Perspectives*, 31(1):183–204.

Kaufmann, K. M., Messner, M., and Solis, A. (2013). Returns to elite higher education in the marriage market: Evidence from Chile. *Available at SSRN 2313369*.

Kirkeboen, L. J., Leuven, E., and Mogstad, M. (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics*, 131(3):1057–1111.

Kleibergen, F. and Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1):97–126.

Liu, H. and Lu, J. (2006). Measuring the degree of assortative mating. *Economics Letters*, 92(3):317–322.

Mansour, H. and McKinnish, T. (2014). Same-occupation spouses: preferences or search costs? *Journal of Population Economics*, pages 1–29.

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

Nielsen, H. S. and Svarer, M. (2009). Educational homogamy how much is opportunities? *Journal of Human Resources*, 44(4):1066–1086.

Petrongolo, B. and Pissarides, C. A. (2001). Looking into the black box: A survey of the matching function. *Journal of Economic literature*, 39(2):390–431.

Schwartz, C. R. (2013). Trends and variation in assortative mating: Causes and consequences. *Annual Review of Sociology*, 39:451–470.

Part I

Appendix

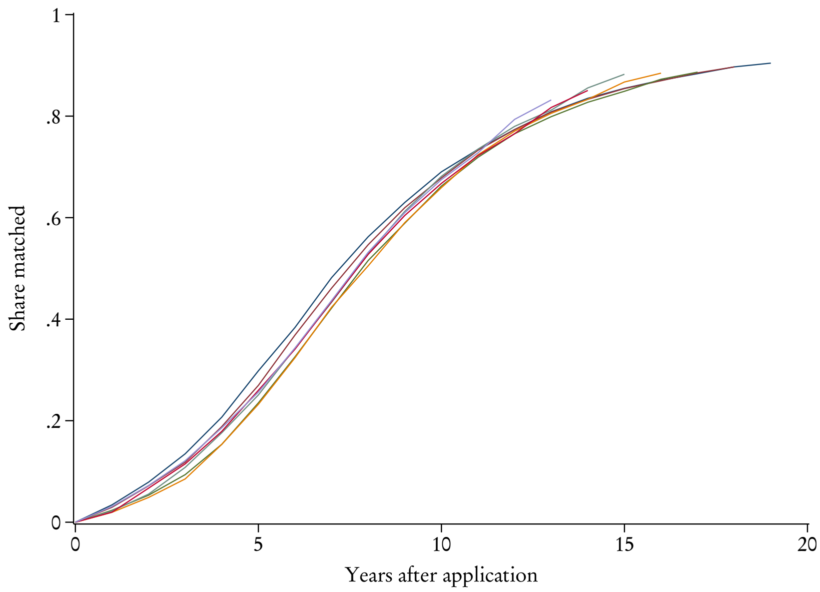


Figure A1. Share matched, by cohort

Table A1. 2SLS homogamy effects, complier shares, contribution to total effect on preferred-field homogamy and assortativity effects by preferred field; average GPA, earnings and homogamy by completed field

	Enroll	First	Complier	Effect		Averages by Field		
	Effect	stage	Share	Share	Assort.	GPA	Earnings	Homog.
Science	0.027	0.224	0.051	0.034	0.056	0.196	0.534	0.069
	(0.020)	(0.019)						
Engineering	0.047	0.214	0.011	0.013	0.141	0.126	0.635	0.050
	(0.043)	(0.038)						
Commerce	0.043	0.148	0.073	0.076	0.047	0.064	0.586	0.118
	(0.038)	(0.012)						
Soc. Science	0.042	0.239	0.080	0.081	0.055	0.312	0.459	0.076
	(0.017)	(0.015)						
Teaching	0.035	0.271	0.128	0.110	0.067	-0.226	0.410	0.051
	(0.013)	(0.012)						
Humanities	0.017	0.198	0.052	0.022	0.024	0.199	0.381	0.100
	(0.026)	(0.017)						
Health	0.031	0.348	0.454	0.347	0.110	-0.169	0.413	0.040
	(0.006)	(0.007)						
Technology	0.038	0.212	0.049	0.045	0.089	0.690	0.713	0.145
	(0.048)	(0.020)						
Law	0.131	0.229	0.057	0.183	0.172	0.449	0.625	0.148
	(0.036)	(0.018)						
Medicine	0.083	0.365	0.045	0.091	0.114	0.798	0.669	0.137
	(0.044)	(0.022)						
Total	0.041							
	(0.006)							

Note: 2SLS estimates of the effect of enrollment on preferred-field-of-study homogamy by field-of-study cf. the specifications outlined in section 7. Complier shares are (relative) sample size weighted first-stage coefficients. Effect shares indicate corresponding relative contributions of the preferred-field-specific effects to the total effect. Assortativity refers to estimates rescaled by the average scaling factor $h^m - h^r$ for treated compliers (i.e. complier y^1) cf. section 6. GPA is standardized in the sample, Earnings are in 100K NOK and measured 13 years after applying. Homogamy refers to observed-field homogamy in the sample. Standard errors are clustered at the individual level.

Table A2. 2SLS homogamy effects, complier shares, contribution to total effect on preferred-institution homogamy and assortativity effects by preferred institution; average GPA, earnings and homogamy by completed institution

	Enroll Effect	First stage	Complier Share	Effect Share	Assort.	Averages by Institution GPA	Earnings	Homog.
HiA	0.158	0.276	0.042	0.069	0.200	-0.112	0.457	0.202
	(0.035)	(0.020)						
HiB	0.091	0.360	0.079	0.074	0.143	0.089	0.464	0.140
	(0.021)	(0.016)						
HiO	0.050	0.384	0.171	0.088	0.094	0.109	0.444	0.080
	(0.010)	(0.011)						
HiS	0.098	0.365	0.073	0.075	0.148	0.026	0.508	0.187
	(0.022)	(0.016)						
HiST	0.117	0.385	0.108	0.131	0.154	0.100	0.470	0.153
	(0.019)	(0.014)						
NHH	0.091	0.282	0.038	0.036	0.102	0.759	0.808	0.147
	(0.040)	(0.021)						
NTNU	0.253	0.231	0.046	0.120	0.324	0.594	0.612	0.275
	(0.056)	(0.018)						
UiB	0.144	0.342	0.046	0.068	0.176	0.480	0.530	0.227
	(0.031)	(0.020)						
UiO	0.118	0.229	0.069	0.084	0.161	0.517	0.509	0.202
	(0.035)	(0.014)						
Other	0.075	0.370	0.328	0.255	0.131	-0.133	0.473	0.090
	(0.008)	(0.008)						
Total	0.095							
	(0.006)							

Note: 2SLS estimates of the effect of enrollment on preferred-institution homogamy by institution cf. the specifications outlined in section 7. Complier shares are (relative) sample size weighted first-stage coefficients. Effect shares indicate corresponding relative contributions of the preferred-institution-specific effects to the total effect. Assortativity refers to estimates rescaled by the average scaling factor $h^m - h^r$ for treated compliers (i.e. complier y^1) cf. section 6. GPA is standardized in the sample, Earnings are in 100K NOK and measured 13 years after applying. Homogamy refers to observed-institution homogamy in the sample. Standard errors are clustered at the individual level.