NBER WORKING PAPER SERIES

PEER INFLUENCE IN COLLEGE APPLICATIONS

Ana M. Gazmuri
Elena Prager

Working Paper 35993
http://www.nber.org/papers/w35993

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2026

We thank Vivek Bhattacharya, Anais Fabre, David Figlio, Chao Fu, JF Houde, Rob Porter, John Singleton, and numerous seminar and conference participants for helpful suggestions. Gazmuri acknowledges funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EURE-0010, grant MATCHINEQ - ANR-22-CE26-0005-01, and The European Research Council (ERC) and the European Union (grant agreement No 101162723— EDDYNCHOICE). Prager gratefully acknowledges funding from the Washington Center for Equitable Growth. We thank Diego Fica for excellent research assistance. We are grateful to Lars Nielsen, Allison Van, and Lily Wang at McMaster University's Secure Empirical Analysis Lab (SEAL) for their support with the OUAC data. The views expressed in this paper are those of the authors and may not represent OUAC, any Ontario university, or any funder. Any errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Ana M. Gazmuri and Elena Prager. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Peer Influence in College Applications
Ana M. Gazmuri and Elena Prager
NBER Working Paper No. 35993
November 2026
JEL No. I20, I23, I24

ABSTRACT

This paper studies how exposure to peers affects college application behavior. We use neighbors who attend different high schools to mitigate concerns about common shocks. Using granular classroom-level data, we find that exposure to peers who have applied to a program increases application rates. Consistent with an information mechanism, there is no additional effect of peers who are admitted. In counterfactuals, the effect is large enough to close urban-rural gaps and gender gaps in applications to selective college programs. We make the (generally NP-hard) application portfolio choice problem tractable using a setting where admission probabilities are observable and conditionally independent.

Ana M. Gazmuri
Universidad Diego Portales
ana.gazmuri@udp.cl

Elena Prager
University of Rochester
Simon Business School
and NBER
eprager@simon.rochester.edu

There are persistent gaps in college applications and major choice by income, geography, gender, and other demographic characteristics (Patnaik et al. 2021; Dynarski et al. 2023). These gaps are not fully explained by student ability or ability to pay, and information frictions play a key role (Hoxby and Turner 2015; Lavecchia et al. 2016; Dynarski et al. 2021). An important potential shifter for teenagers making college application decisions is their peers. Peers shift the extensive margin of any college application and attendance (Bobonis and Finan 2009; Mora and Oreopoulos 2011; Abramitzky et al. 2021), but little is known about their effects on the intensive margin of where to apply. This paper studies how much changing the composition of peers can narrow the gaps in college application decisions, especially gaps in applications to STEM and selective programs.

Estimating the effect of peer exposure on application decisions requires overcoming the usual identification challenges in studies of peer effects. We leverage an institutional setting, the Canadian province of Ontario, that helps with these challenges. We use it to study how exposure to peers who applied, were admitted, or matriculated to a given program shifts students' decisions about which programs to apply to. The first identification challenge is a reflection problem: when two peers make similar choices, it is difficult to identify which peer influenced the other. There are also correlated unobservables: peers may make similar choices because they were exposed to a common unobserved shock. For example, a common teacher might encourage both peers to apply to her alma mater.

To mitigate the reflection problem, we take advantage of Ontario's high school curriculum resulting in students from different high school cohorts sharing a classroom. We study the effect of within-classroom exposure to older cohorts who applied to, were admitted to, or matriculated to a given program on the application decisions of subsequent cohorts. Granular data on which students within a school also share a classroom reduce false positives in the exposure measure. The cross-cohort strategy mitigates concerns about the older cohorts being influenced by the younger ones. However, the problem of common unobserved shocks persists even across cohorts. Our identification strategy therefore uses a second layer: within-neighborhood exposure to students who live nearby but attend a different high school. In Ontario, students living in the same granularly defined neighborhood—roughly equivalent to a US Census block group—may attend different high schools. We therefore use *neighbors'* exposures to older cohort schoolmates who have a history with a given

program.

In addition to the challenges of identifying peer effects, we face the challenge that observed college application choices are not directly informative about preferences. This is because applications are also a function of (typically unobservable) admission probabilities. Students in Ontario apply directly into majors within colleges, typically applying to three or four total programs. Failure to apply to a competitive program does not necessarily imply that the student does not want to attend that program; the student may simply correctly anticipate that she will not be admitted. Moreover, the decision to apply to a given program is also a function of preferences over other programs and the full set of admission probabilities. This challenge is common to settings where an agent must assemble a portfolio with uncertain returns, or, in the case of college applications, uncertain admission probabilities.

The structure of college admissions in Ontario helps to resolve the difficulty of estimating preferences from observed application choices. During our sample period, Ontario's college admissions are based almost exclusively on students' high school transcripts (Card and Payne 2021), which we observe.[1] This feature makes admissions probabilities effectively observable to the econometrician, allowing us to account for college applications' dependence on those probabilities. This allows us to make students' application portfolio problem tractable by recasting it as a modified discrete choice problem. Ordinarily, solving for preferences from application choices involves an intractable combinatorial problem (even with known admission probabilities), as described later in this introduction's discussion of counterfactuals. Our estimator leverages the fact that, after conditioning on observable admission probabilities, application behavior is informative about preferences (Beuermann et al. 2023). If a student has an equal probability of being admitted to two programs but applies only to the first program, our estimator infers that the student prefers the first program to the second. This allows us to estimate the utility parameters without solving the entire portfolio problem in each iteration. This estimator hinges on students observing their own admission probabilities. In Ontario, admissions criteria are easily observable to students despite being major-specific. Indeed, we show in the data that students' response to their admission probabilities is consistent with observability and rationality, such as stronger students applying to more

[1]Other factors commonly used in US college admissions are almost wholly absent, including standardized tests, application essays, interviews, letters of recommendation, extracurriculars, and high school-specific admissions thresholds.

selective programs.

These institutional features allow us to estimate how application choices respond to exposure to (neighbors') peers who have a history with a given program. Our approach relies on a two-stage structural model of college decisions. In the first stage, a student chooses multiple college programs to apply to based on her underlying preferences and the probability of being admitted. In the second stage, she receives admission offers and chooses one program to attend based on her underlying preferences. Those underlying preferences are potentially shifted by exposure to peers who applied to, were admitted to, or ultimately attended a program. To deal with the reflection and common shock problems discussed above, we only use exposure to the older-cohort classmates of neighbors who attend a different high school. The structural approach allows us to account for interdependencies across programs. For example, a student who already has "safe" programs in her application portfolio is more likely to add a "reach" program. This interdependence would not be adequately captured by a regression that considered a student's decision about each program separately, and is not addressed by approaches developed for settings where only one alternative is chosen, such as De Giorgi et al. (2010).

We find that exposure to peers plays an important role. Exposure to a peer who applied to a given program increases the probability of application. The magnitude is comparable to a 0.4 standard deviation reduction in distance from home, in a setting where many students enroll near home. The effect is driven entirely by same-gender peers, consistent with known effects of peer gender composition on major choice (Anelli and Peri 2019; Zölitz and Feld 2021). Whether the peer is admitted to a program has no additional effect on the focal student's choice. This is consistent with the peer exposure operating through information or awareness of programs, rather than with students updating their beliefs about admission probabilities. Our results do not rule out an additional role for insider information about programs. While we find no such effect using our preferred measure of within-classroom exposure, we find an effect of a student matriculating at a program using a second measure of exposure (that has lower false negatives but higher false positives). We interpret this result as a possible role for insider information; it is unlikely to be explained by a preference for programs where the student has friends or social connections, because the peer exposure is to neighbors' classmates the student is unlikely to know directly.

Using these estimates, we conduct counterfactual simulations to evaluate the effects of changes

to students' peer exposure. This requires solving for students' new optimal application portfolios under each counterfactual scenario. Ordinarily, this would require evaluating millions of possible combinations of programs for each error draw for each student. Even with known admission probabilities, this is an NP-hard problem that cannot be solved in polynomial time. We leverage Ontario's admissions rules to make this problem tractable. In the special case where a student's admission probabilities are known and independent across programs, there exists a tractable solution to the NP-hard portfolio problem (Chade and Smith 2006). The requirements for this special case are difficult to satisfy. For example, they rule out any inputs to admissions decisions that can unobservably raise a student's admission probabilities across multiple programs, such as strong interview skills. Ontario satisfies this requirement because admission probabilities are based almost solely on observable transcript information. We show in the data that the student's admission outcomes at other programs and the identity of a student's high school are not predictive of admissions.

Our first counterfactual evaluates whether peer exposure can narrow gender gaps in applications. In the data, high-ability women apply to less selective programs than high-ability men, even after conditioning on granular high school GPA and performance in STEM courses. We find that exposing women to same-gender peers who applied to selective programs would close the gender gap in applications among the top decile of students, and would result in women applying more aggressively than men at every decile below that.[2] A back-of-the-envelope calculation suggests that the shift in applications would also close 26 to 29 percent of the gap in program completion rates between high-ability women and high-ability men, assuming that only 20 percent of variation in college completion rates is driven by causal effects of colleges. Results are similar for selective STEM programs. We conduct a similar exercise to study the gap between urban and rural students. Students outside large cities apply to less selective programs, partly because selective programs tend to be located farther from home (Fabre 2026). We find that exposing these students to peers who applied to selective programs is equally effective for closing the urban-rural gap as eliminating differences in the spatial distribution of students. The importance of peer exposure is illustrated by its ability to generate the same shifts as a much more invasive intervention, relocation.

Our data do not allow us to study downstream outcomes such as college persistence and post-college earnings. We therefore cannot evaluate the long-run effects of applying to more selective

[2]The large effect is explained by the fact that endowing a student with exposure to *every* selective program is an infeasibly large intervention.

programs. Nevertheless, it is unlikely that our counterfactuals push students to attend programs where they are not equipped to succeed. The optimal portfolio algorithm prevents students from applying to programs where they have little chance of being admitted, and in the counterfactuals, we keep admissions thresholds fixed. A potential drawback of fixing the admissions thresholds is that it assumes no capacity constraints. Therefore, we bound the role of flexible capacity using a counterfactual with additional structure. We find that allowing admissions thresholds to adjust while imposing capacity constraints is unlikely to change the qualitative conclusions. If anything, our primary results may understate the effectiveness of peer exposure on narrowing demographic gaps under fixed capacities.

Our results may help to reconcile the ambiguous effectiveness of low-intensity treatments on decisions about college. There is ample evidence that peers causally influence whether students apply to and attend college (Bobonis and Finan 2009; Mora and Oreopoulos 2011; Abramitzky et al. 2021; Barrios-Fernández 2022). There is also some evidence that very close peers such as siblings influence which colleges students apply to, and mixed evidence on whether they influence major choice (Altmejd et al. 2021; Aguirre and Matta 2021; Avdeev et al. 2024). However, there is little evidence on whether low-intensity peer exposures—of which there are many more than exposures to siblings or close friends—influence decisions about which colleges or majors to apply to, conditional on any college. (The two notable exceptions are De Giorgi et al. (2010), who find effects of classmates on major choice; and Avdeev et al. (2024), who find effects of neighbors.) Such decisions are less drastic than whether to apply or attend at all, and may therefore be more responsive to low-intensity treatments. Our results suggest this is indeed the case. This is also consistent with the success of other low-intensity interventions at shifting which colleges students apply to (Hoxby and Avery 2013), even as they have generally shown less promise than costlier interventions for shifting the extensive margin of college application and attendance (Lavecchia et al. 2016; Dynarski et al. 2023).

Our findings of an effect of peer applications, but no additional effect of peer admission or matriculation, are consistent with the peer influence operating through an information mechanism. This is distinct from the aforementioned literature on major and college choices. In those papers, the identifying variation typically does not allow the authors to study the effect of a peer merely applying (Altmejd et al. 2021; Avdeev et al. 2024; De Giorgi et al. 2010). The existing literature

is therefore also consistent with students choosing programs because of a preference for having friends or mentors in one's program, or having access to insider information about the program. In our setting, an application by a neighbor's classmate, whom the focal student is unlikely to know directly, shifts application decisions. This suggests at least some role for awareness rather than preferences. That we find effects via these loose social ties also suggests one mechanism through which socioeconomic diversity benefits young people from low socioeconomic status backgrounds (Chetty et al. 2022). Even low-intensity exposure to peers who apply to high-earnings potential colleges or majors may shift the exposed student toward higher-return programs.

The remainder of the paper is organized as follows. Section 1 describes the pertinent institutional features of Ontario. Section 2 describes the data. Section 3 presents the model of college application and matriculation. Section 4 presents the estimates of peer influence and conducts counterfactual simulations.

1. College Applications in Ontario

The majority of college-bound high school students in Ontario stay within Ontario for their post-secondary education. There are 21 universities in Ontario during our sample period.[3] These are all public universities and are subject to Ontario's rules governing admissions and tuition. Tuition sticker prices are appreciably lower than in the US, and vary little across programs due to regulation (Appendix B.2 and Figure A.2). The logistics of the application process are handled through a single central clearinghouse, the Ontario Universities' Application Centre (OUAC). Appendix B.1 describes the role of OUAC in the application process.

Ontario law shrinks the typical gaps between what is observable to universities, students, and the econometrician. Universities' admission decisions depend almost exclusively on a student's high school courses and grades. Since we observe both, the variables used to determine admission are observable to the econometrician. Along with observing admissions decisions, this allows us to construct reliable admission probabilities for any student applying to any program.

Admissions decisions are based almost solely on performance in grade 12 courses. Several factors that are important determinants of admission to US colleges are wholly absent from the admission

[3] In Ontario, post-secondary institutions offering four-year degrees are called "universities". The word "college" is used to describe community colleges and similar institutions offering associate's degrees. In the paper, we instead use "university" and "college" interchangeably to mean a four-year degree-granting institution.

process in Ontario (Card and Payne 2021). There is no standardized testing.[4] Universities are also prohibited from considering the identity or characteristics of a student's high school in the admissions decision.[5] They are required to treat numerically equivalent grades from any two accredited high schools as identical. Supplementary materials such as application essays, portfolios, extracurriculars, and letters of recommendation are very rarely required or considered. The key exception is for admission to performing and fine arts programs, where students typically submit a portfolio or complete an audition; these factors are unobservable to the econometrician.[6] Figure A.1 shows that, although there is a steep gradient in admission probability with respect to high school GPA (first panel), the gradient is flat with respect to income after conditioning on high school GPA (second panel), consistent with admissions being based almost solely on high school performance.

Admissions requirements vary by both university and major or program of study. This is empirically important because the overwhelming majority of students (94 percent of applications in our data) apply directly into a specific major.[7] Each program specifies a set of required and sometimes recommended grade 12 courses. For example, engineering programs typically require calculus, one or two additional math courses, physics, and sometimes chemistry. Pre-med and life sciences programs typically require one or two math courses, biology, and chemistry. Humanities programs often do not require any particular grade 12 courses beyond the already-compulsory English course (Appendix B.3 describes the high school curriculum). Admission is typically based on the GPA from the student's best six grade 12 courses, including the required and recommended courses for that program.

Which grade 12 courses a student completes therefore determines which undergraduate programs are likely to admit her. We take as given students' high school courses, and study the downstream decision of college applications conditional on high school courses and grades. However, course enrollments provide crucial identifying variation in the set of peers a student is exposed

[4]International applicants are required or encouraged by some universities to submit TOEFL English scores and/or other standardized test scores, but these requirements do not apply to students applying from within Ontario.

[5]Disparities across high schools are also less pronounced than in the US, as described in Appendix B.4.

[6]Toward the end of our sample period, a few universities' business programs began to use students' extracurricular activities and application essays as part of the admission process. These are also unobservable to the econometrician.

[7]Switching majors after matriculation is substantially rarer in than in the U.S. Switching typically requires the student to apply internally, and the application is only approved if the student meets performance thresholds in courses required for the new desired major. It is typically difficult to switch from a less competitive major with loose initial admission criteria to a more competitive major.

to, even conditional on taking the same courses (Section 2.3).

There is a fee for submitting applications. The base fee ranges from CAD $115 to $140 during our sample period, and allows a student to submit up to three application. Students can pay an additional piece-rate fee of CAD $38 to $47 for each application beyond the third. Most students apply to only a few programs. The modal number of applications per student is three, the median is four, and the 90th percentile is seven. Students may apply to more than one major within the same university, in which case those applications are typically evaluated separately.

Students can learn admission requirements by consulting eINFO, OUAC's comprehensive publication that collects information about every program in one place and presents the information in a standardized format accessible online or in hard copy. Appendix B.1 describes the eINFO booklets in more detail. Figure 1 shows an example set of admission requirements. Students observe the set of required and recommended courses for admission and the approximate high school GPA threshold for admission in the previous application cycle.[8]

Figure 1: Admission requirements example

OUAC Code	Program Name	Previous No. Enrolled	Previous Grade Ranges	Ontario Secondary School Prerequisites
GBK	Accounting		76-81%	ENG4U MHF4U One additional 4U math Three additional 4U/M courses
GAD	Adult Development		76-80%	ENG4U 4U math One of SBI4U or SCH4U Three additional 4U/M courses SBI4U is strongly recommended
GPA	Agriculture		76-80%	ENG4U MHF4U Two from SBI4U, SCH4U, SPH4U Two additional 4U/M courses
GS	Applied Human Nutrition		81-85%	ENG4U SBI4U SCH4U One 4U math Two additional U/M courses You must take both SBI4U and SCH4U.

Note: Screenshot from one university's section of the 2015 OUAC eINFO booklet, the key information source for high school students about available undergraduate programs published by OUAC. The three-letter OUAC codes identify the programs that accept separate applications. The Previous Grade Ranges column reports the approximate high school GPA required for admission to the program in the previous application cycle. The alphanumeric codes in the Prerequisites column are standardized codes for Ontario high school courses.

In the Ontario context, both the student and the econometrician observe nearly the entirety

[8]High school grades are assigned on a 100-point scale. Grade inflation is at most moderate during our sample period: grades close to 100 are rare, and fewer than half a dozen programs have minimum admission GPAs above 90. Appendix B.3 describes the grading scheme in more detail.

of the information used by admissions offices. The granular admissions requirements and previous year's GPA published in eINFO, along with the absence of other inputs into admissions decisions, allow students to estimate their admissions probabilities far more accurately than for applications to selective US colleges. The minimum GPA required for admission is quite stable within program over time. Remaining uncertainty about admission may arise from application deadlines preceding final course grades, small shifts in admission thresholds over time, and being close to the admission threshold. These features allow us to treat admissions probabilities as observable ex ante, even if the realization of the admission outcome is uncertain. Section 3.5 describes how we construct admission probabilities and shows that the constructed probabilities meaningfully predict student behavior.

2. Data and Definitions

2.1 Data Description

We use microdata from the OUAC application clearinghouse on the universe of all Ontario secondary school students applying to Ontario universities from 2011 to 2016. The data include basic applicant demographics, academic performance in high school, the list of programs to which the student applied, admissions decisions for those programs, and the student's matriculation decision.

Ontario is Canada's largest province, with a population of approximately fourteen million, which is nearly 40 percent of Canada's total. We observe approximately 90,000 applicants per year for a total of approximately 546,000 applicants. Slightly more than half of applicants (54.5 percent) are women, consistent with women pursuing college degrees at higher rates than men. A substantial minority of applicants (15.5 percent) are foreign-born immigrants (Figure A.3), reflecting Ontario's large immigrant population; 29 percent of Ontario's population is foreign-born (Census 2011, Census 2016).[9] Immigrants and non-immigrants apply to similarly competitive programs. Admission rates to undergraduate programs are high. More than three-quarters of applicants in the data are admitted to at least one program. Table A.1 summarizes application patterns at the level of the high school; these figures overweight small schools relative to a mean across all students.

The OUAC data contain information on each key aspect of applications: the student's final

[9] These counts understate the true proportion of applicants who are immigrants, because many naturalized citizens are not coded as immigrants in the data.

year academic performance in each high school course, the set of universities and majors to which the student applies, the admissions decisions for those applications, and the program to which the student chooses to matriculate.[10] The course data allow us to infer which pairs of students interacted with schoolmates across cohorts by taking the same course in the same semester, creating within-high school variation in the information environment. The data also contain information about the student's high school, age and sex, home neighborhood, average income by neighborhood, and noisy variables measuring immigrant status and mother tongue.

While these data are extremely detailed, they are missing some pertinent information. We do not observe dates of the university's admission decision nor of the student's matriculation decision. The data contain only the students' final grades in their courses, whereas admissions may be made based on interim grades for courses that are still in progress. Nevertheless, the large representative sample and the observability of all key variables determining admissions are key strengths of our data and setting.

2.2 Majors of Study

The majority of applicants in Ontario (94 percent) declare a major at the time of application. The number of separate programs, approximately one thousand, is too large for tractable estimation. We therefore classify undergraduate programs into eighteen categories of majors according to the field of study and admission requirements. In the empirical analysis, a "program" is a major-university pair. Appendix C.1 describes how we aggregate programs into the eighteen major categories, and Table C.1 lists the fraction of total applications made to each category. The largest categories in terms of applications are business, commerce, and accounting at 19 percent of applications; engineering and computer science at 15 percent; and biology, life sciences, and pre-med at 13 percent.

2.3 Peer Groups

Our identification strategy relies on exposures of students across cohorts that can lead to across-cohort information diffusion, and further diffusion through interacting with neighbors outside the

[10] One university, the Royal Military College of Canada (RMC), is not included in the OUAC data because the military college uses a separate admissions system. RMC enrolls approximately 250 new full-time undergraduates per year, less than 0.5 percent of our sample size.

high school. To study the effect of peer exposure on application decisions, we construct peer groups using information on whether a pair of students was enrolled in classes together. The high school curriculum in Ontario affords students a substantial degree of choice among courses (Appendix B.3). As a result, two students in the same or adjacent cohorts of the same high school may sit in few or even no classrooms together throughout their high school career. Large high schools also generate variation in students' degree of overlap even after conditioning on similar course loads.[11] This provides variation in our measure of exposure to peers in older cohorts who apply, are admitted, or matriculate to a given program.

Given that students in older cohorts apply to college before younger cohorts, this strategy mitigates reverse causality concerns. Nevertheless, one may worry that even across-cohort correlation in application behaviors could be explained by common unobservables within a high school. To deal with common shocks within a high school, we take advantage of students who live in the same neighborhood but attend different high schools. This can happen due to a combination of magnet programs and parallel tracks, described in detail in Appendix B.5. For each focal student, identification relies on the student's *neighbors'* exposure to older peers in the neighbor's high school that the focal student does not attend.

Students may enroll in courses that are formally designated to be a different grade level. For example, a student who completes grade 11 biology in the first semester of grade 11 may then enroll in grade 12 biology during the second semester of the same academic year. Students can therefore share classrooms with classmates from other cohorts. We infer whether a student shared a classroom with each student from her high school in the preceding two cohorts from the course enrollment data.[12]

We use two measures of within-classroom peer exposure. The primary measure is a binary measure of whether two students were in the same low-enrollment course in the same semester. We use this measure to deal with noisy classroom information in the data. Rather than observing

[11] High schools vary dramatically in size. The largest schools tend to be in dense neighborhoods of urban areas. The median school has approximately 170 students per cohort; the 10th percentile has approximately 20, and the 90th percentile has approximately 330. Half of students attend schools with cohorts of size 250 or larger. In large schools, two students taking the same course are likely to be in separate course sections. In small schools, only one section of a given course may be offered per semester or even per year, particularly if it is an elective. In a small school, two students both aiming to apply to engineering programs may take all their math and science courses together in the same sections, but this is unlikely in a large school.

[12] A two-year cutoff corresponds to a student being in the same class during her freshman year of high school as a high school junior. In robustness checks, we find no effect of cohorts earlier than the two preceding cohorts. We therefore use the more parsimonious specification.

classroom identifiers, we observe the detailed course code (e.g. EWC4U is writer's craft, a grade 12 English elective course meant to prepare students for university-level coursework), the end date of the course, the mode of delivery (regular daytime instruction, summer school, or night school), the language of instruction (typically English or French), and the type of instruction (i.e. regular course or co-op course). We say that the students share a classroom if they are in the same such cell, and the cell has at most ten students. This conservative measure allows us to be confident that for the overwhelming majority of these cells, there was a single course section taught in one physical classroom.

The small classroom measure discards potentially useful information from courses with larger enrollments. As a robustness test, we therefore use a second measure that calculates continuous probabilities of two students interacting in at least one classroom. This measure reduces false negatives at the cost of raising false positives. For inferred classrooms with enrollments of 25 or fewer, we assume that there is only one physical classroom per cell. For inferred classrooms with enrollments above 25, we assume that there may be more than one physical classroom per cell; this is most common in large schools for common courses such as English and calculus. In such cases, we assume that the probability of two students i, i' actually being in the same physical classroom is given by a continuous probability decreasing in N_c, the number of students in the inferred classroom, $\Pr\{i, i' \in \text{same classroom } c\} = 25/N_c$. For example, two students in an inferred classroom with an enrollment of 50 have an inferred probability of being in the same physical classroom of 0.5.

The final step is to construct the total probability of being in at least one physical classroom together for each pair of students. For two students sharing $c \in \{1, 2, \ldots, k\}$ inferred classrooms, we define the total probability that they shared at least one physical classroom as:

$$\begin{aligned}\Pr\left\{i, i' \in \text{at least one same classroom}\right\} =& 1 - \prod_{c=1}^{k} \left(1 - \Pr\left\{i, i' \in \text{same classroom } c\right\}\right) \\ =& 1 - \prod_{c=1}^{k} \left(1 - 25/N_c\right)\end{aligned}$$

2.4 Suggestive Evidence of Peer Effects

We begin with descriptive regressions consistent with student application behavior being responsive to peers' application history. For each student-program pair, we regress an indicator for applying

to that program on whether the student is exposed (through a neighbor) to peers who previously applied. We use exposure through a close neighbor attending a different school, rather than within-school exposure, to deal with correlated unobservables within school (Section 3.4).

Table 1 shows the results. The baseline probability of application to a given program is 1.1 percent, and the probability is 2.6 percentage points higher if exposed to peers who applied in the prior year (Column 1). On average, students are less likely to apply to more selective programs, but this does not dampen the coefficient on peer exposure (Column 2). Flexibly controlling for students' application portfolio size using student fixed effects does not meaningfully affect the results (Column 3). The positive relationship between peer exposure and applying is larger when the exposure is from a same-gender peer (Column 4).

Table 1: Correlations between Applications and Peer Connections

	Probability of Applying			
Neighbor interaction probability (t-1)	0.0258***	0.0245***	0.0303***	
	(0.0002)	(0.000160)	(0.0002)	
Program's Median Admission Probability		-0.0138***	-0.0136***	-0.0134***
		(0.0001)	(0.0001)	(0.0001)
Neighbor interaction probability (t-1) (same gender)				0.0301***
				(0.0002)
Neighbor interaction probability (t-1) (different gender)				0.0170***
				(0.0002)
Constant	0.0111***	0.0202***	0.0199***	0.0196***
	(0.00003)	(0.0001)	(0.0001)	(0.0001)
N	17629264	17,629,264	17,629,264	17629264
Student Fixed-Effects			X	X

Notes: Linear probability model where the dependent variable is an indicator for student i applying to a program j

These results are suggestive of the possibility that students transmit information to the peers with whom they interact during high school, especially same-gender peers. However, conclusions from these regressions are limited by the regressions' treatment of a student's decision to apply to a given program as independent of her decision to apply to any other program. In reality, application portfolios are strategic: a student should optimally apply to some "reach" programs, but also to some programs where she is likely to be admitted. A positive correlation between peer exposure and applying could be driven by differential exposure for programs that are already a good strategic addition to a student's application portfolio, because these regressions do not control for the other programs a student applies to. Finally, because application portfolios are limited in size, applying to one program can mechanically reduce the probability of applying to another. We next turn to a

structural model that addresses these shortcomings by jointly estimating applications and program choices.

3. Structural Model and Estimation

This section describes the structural model we use to jointly estimate students' application choices and subsequent decisions about which program to attend, among those where the student is admitted. Section 3.1 builds up to the empirical specification, starting with the theoretical underpinning of college application portfolios. Section 3.2 then describes the discrete choice problem faced by the student in deciding among offers of admission. Section 3.3 describes how we parameterize both stages in order to take the model to the data. Finally, Section 3.4 describes the identification strategy based on neighbors' peer exposures.

3.1 Optimal Application Bundles

A student makes two sequential decisions. The first, which occurs in the winter of the student's final year of high school, is the decision of which programs to submit applications to. The second, which occurs in the spring of the same school year, is the decision of which program to attend (if any) among those to which the student was admitted (if any).

We begin with a discussion of the student's first decision: which programs to apply to. This is a portfolio choice problem. Call the utility from attending a program its ex post utility, and the expected utility at the application stage the ex ante utility. The student may know her ex post utility of attending each program, but she faces uncertainty over which programs will admit her. The portfolio of applications therefore maximizes the student's ex ante expected utility, which is a function of both how much she likes each program and how likely she is to be admitted.

In general, solving for a student's optimal application bundle is NP-hard (Chade and Smith 2006). A student may not apply to the programs she most prefers if she is unlikely to be admitted, making it difficult to infer preference parameters from observed application decisions. We make the estimation tractable by recasting the portfolio problem as a modified discrete choice problem. Our estimator relies on the insight that after conditioning on admission probabilities, observed application choices are informative about true preferences. Suppose a student has equal probabilities of admission to programs s and s', but only includes s in her application portfolio. Then it

must be that s has higher ex post utility than s' (i.e., s is preferred to s'). A formal proof of this property of the optimal application portfolio is given by Chade and Smith (2006). Therefore, if the econometrician can observe the admission probabilities and appropriately condition on them, then a discrete choice estimator can recover true preferences.

We therefore follow Beuermann et al. (2023) in constructing a maximum likelihood estimator that conditions on admission probabilities in order to recover truthful preferences over programs. In this formulation, the student chooses programs to include in the application portfolio by maximizing a flexible function of a program's ex post utility and admission probability, rather than ex post utility alone. Denote by $\delta_{is} + \varepsilon_{is}$ the student's ex ante utility of program s, where ε_{is} is an i.i.d. Type I extreme value idiosyncratic error term. δ_{is} is a function of admission probability, program characteristics, and exposure to (neighbors') peers who have applied, been admitted, or matriculated to the program. In the empirics, we will parameterize this flexibly as described on page 21. Observe that for any pair of programs s and s', if s is included in the student's application bundle and s' is not, then it must be that $\delta_{is} + \varepsilon_{is} > \delta_{is'} + \varepsilon_{is'}$. This gives us an incomplete ex ante ranking over programs for each student: every program observed to be in the student's application bundle B_i must have a higher ex ante utility than every program excluded from the bundle. Moreover, after conditioning on admission probabilities, programs in the bundle must have a higher ex post utility than programs excluded from the bundle.

The probability of observing a given application bundle B_i can be expressed as:

$$\Pr\left(B_i\right) = \Pr\left(\delta_{is} + \varepsilon_{is} > \delta_{is'} + \varepsilon_{is'} \quad \forall s \in B_i, s' \in \mathcal{S} \backslash B_i\right)$$

If the student's truthful ranking of programs in terms of ex post utility were observable, then this expression would become an exploded logit:

$$\begin{aligned}\Pr\left(B_i\right) = & \frac{\exp\left(\delta_{i1}\right)}{\sum\limits_{s \in B_i} \exp\left(\delta_{is}\right) + \sum\limits_{s' \in \mathcal{S} \backslash B_i} \exp\left(\delta_{is'}\right)} \cdot \frac{\exp\left(\delta_{i2}\right)}{\sum\limits_{s \in B_i \backslash \{1\}} \exp\left(\delta_{is}\right) + \sum\limits_{s' \in \mathcal{S} \backslash B_i} \exp\left(\delta_{is'}\right)} \\ & \cdot \ldots \cdot \frac{\exp\left(\delta_{i\|B_i\|-1}\right)}{\sum\limits_{s \in B_i \backslash \{1,2,\ldots,\|B_i\|-2\}} \exp\left(\delta_{is}\right) + \sum\limits_{s' \in \mathcal{S} \backslash B_i} \exp\left(\delta_{is'}\right)} \cdot \frac{\exp\left(\delta_{i\|B_i\|}\right)}{\exp\left(\delta_{i\|B_i\|}\right) + \sum\limits_{s' \in \mathcal{S} \backslash B_i} \exp\left(\delta_{is'}\right)}\end{aligned}$$

where $\|B_i\|$ is the size of the optimal application bundle.

Chade and Smith (2006) show that in settings where a student's admission probabilities are

independent across programs, there exists a tractable solution to the optimal application portfolio. The validity of our estimator therefore rests on the measurability and conditional independence of admission probabilities within a student across programs, after conditioning on student observables. For example, if the admission probabilities depended on unobservables, then unobservably weak students would apply to less selective colleges, biasing downward the estimated preference for selectivity. If the admission probabilities were correlated within a student across programs, then the assumption of i.i.d. error terms would be violated, also inducing bias. Ontario's unusual admissions environment is therefore key to making our estimation tractable. We formally test the institutionally motivated conditional independence assumption in Section 3.5. The conditional independence assumption is more appropriate in the Ontario context than in some contexts where it has also been used, but where unobservables play a larger role in admissions (Fu et al. 2022).

Rankings within the bundle. We observe only whether a program was included in the bundle, not the student's truthful rankings within the bundle. In the absence of an incentive-compatible central assignment mechanism, nothing in our setting compels students to truthfully rank the programs to which they apply.[13]

We make a simplifying assumption: that the programs in the bundle are equally preferred to one another for the purposes of building the application bundle. Then the probability of an application bundle becomes:

$$\Pr\left(B_i\right) = \prod_{s \in B_i} \frac{\exp\left(\delta_{is}\right)}{\exp\left(\delta_{is}\right) + \sum_{s' \in \mathcal{S} \setminus B_i} \exp\left(\delta_{is'}\right)} \tag{1}$$

An alternative would be to integrate over all possible permutations of preference rankings within the bundle. This is computationally costly and gains little because, in our setting, it is approximately equivalent to estimation with the assumption of equal preferences. It can be formally shown that Equation 1 is an approximation of the full probability expression that incorporates unequal preferences over programs within the bundle (Breslow 1974). The approximation performs well if the size of the application bundle is not large relative to the number of all possible programs

[13]Students indicate an order of programs when submitting their applications through the OUAC clearinghouse, which OUAC calls a "ranking," but the data are not consistent with this being a preference order. Among students admitted to more than one program, 33.5 percent choose to matriculate at a program that their ordering would suggest was less preferred than another program to which they were also admitted. All universities state that they do not consider how the student orders them relative to other universities in the application portfolio (OUAC 2016). The fourth row of Appendix Table C.2 tests this formally: we find little evidence of universities' admission probabilities varying with the reported ordering.

(Allison and Christakis 1994). In our setting, application bundles are much smaller than the number of options. Students face a choice among more than 300 programs. The median student applies to just four programs, and the 90th percentile student applies to seven programs (Section 1).

Equation 1 discards information from the size of the observed application bundle. In principle, the observed bundle size implies that the most preferred program the student excludes from the bundle contributes less to the expected payoff from the bundle than the marginal cost of an additional application. Because application costs are negligible (Section 1), the size of observed application bundles is not very informative in practice, and we simply omit application costs from the estimation. This also substantially reduces computational time, which is a limiting factor in our setting because we are required to use a specific restricted server with limited computational capacity. Moreover, so long as a student's marginal cost of an additional application is fixed across programs, that application cost will not contribute variation to the estimation, because identifying variation must come from across alternatives within a choice set.

3.2 Matriculation Decisions

In spring of their application year, students learn their admissions outcomes. They then make a choice among the programs to which they have been admitted (if any) and the outside option. This is a standard discrete choice problem, with each student choosing one program at which to matriculate. Unlike in the US where students can apply through the binding early decision process, there is no dynamic component to Ontario students' choices. Ontario universities may make their admissions decisions on a rolling basis, but they are prohibited from asking students to accept or decline offers before a fixed date that comes after universities must issue admissions decisions. The earliest universities can require students to decide among offers of admission is typically the first Monday of June, and universities must typically communicate admissions decisions to students by late May.

We model a student's matriculation decision as a maximization of the student's ex post utility, rather than the ex ante utility used in Section 3.1. The choice set is defined as the set of programs to which the student was admitted (plus an outside option), so the relevant ex post admission probabilities within the choice set are equal to one.

Denote the ex post utility of a program as $\nu_{is}+\eta_{is}$, where ν_{is} is the systematic component and η_{is}

is an i.i.d. Type I extreme value idiosyncratic shock. ν_{is} is a function of program characteristics and exposure to (neighbors') peers who have applied, been admitted, or matriculated to the program. Denote by $A_i \subseteq B_i$ the set of programs student i is admitted to, among the bundle B_i to which she applied. The probability that she chooses to enroll in program $t \in A_i$ is given by

$$\Pr(t|A_i) = \frac{\exp(\nu_{it})}{\sum_{s \in A_i} \exp(\nu_{is}) + \exp(\nu_{i0})} \tag{2}$$

where ν_{i0} is the mean utility of the outside option.

The outside option includes both the outcome of not enrolling in a four-year undergraduate program and the outcome of enrolling in an undergraduate program outside of the province of Ontario. In the data, we do not observe which of these two outcomes occurs. We allow the value of the outside option to differ by the quality of the student and the neighborhood's average income. This captures the fact that high-ability and/or high-income students observed to choose the outside option are disproportionately likely to attend an out-of-province program that we do not observe, rather than not attending any program. In our data, 28 percent of admitted students choose the outside option.

3.3 Parametrization and Estimation

We combine information from the observed application bundles (Equation 1) and observed matriculation decisions (Equation 2) to jointly estimate all utility parameters using maximum likelihood.

The total likelihood for student i is then:

$$\Pr(B_i)\Pr(t|A_i) = \left[\prod_{s \in B_i} \frac{\exp(\delta_{is})}{\exp(\delta_{is}) + \sum_{s' \in \mathcal{S} \setminus B_i} \exp(\delta_{is'})}\right] \left[\frac{\exp(\nu_{it})}{\sum_{s \in A_i} \exp(\nu_{is}) + \exp(\nu_{i0})}\right] \tag{3}$$

This makes the assumption that the idiosyncratic error term at the ex post stage η_{is} is independent from the idiosyncratic error term at the ex ante stage ε_{is}. We relax this assumption below, after discussing how δ_{is} and ν_{is} are parameterized for estimation.

Parameterizing ex post utility. Recall that we aim to estimate the effect of peers on students' application decisions. Of course, students' preferences over universities and majors are not only a function of their peers. Some program characteristics may be vertical, such as the prestige of

the university. Others may be horizontal. For example, students may prefer to attend a university closer to home, as is common in Ontario. Students who are quantitatively inclined may prefer engineering or math programs to humanities programs.

In the estimation, we parameterize student i's ex post utility from program s as a function of inherent program characteristics and student characteristics, denoted by X_{is}; and the student's (neighbors') peers' history with the program, denoted by H_{is}. Both X_{is} and H_{is} are observable; we use separate notation for them solely to clarify the exposition. The vector X_{is} includes major fixed effects and university fixed effects.[14] The vector H_{is} encodes the application, admission, and matriculation history of student i's (neighbors') peers to program s. The ex post utility $\nu_{is} + \eta_{is}$ is then given by:

$$\nu_{is} + \eta_{is} = \beta_{1k(i)} X_{is} + \beta_{2k(i)} H_{is} + \eta_{is} \tag{4}$$

The subscript $k(i)$ denotes a student type, defined below on page 21. The term $\beta_{1k(i)}$ measures student valuation for a given set of program characteristics, in the absence of any peer history with the program. The term $\beta_{2k(i)}$ measures student valuation for peers applying to, being admitted to, or attending a program. For example, students may prefer a program in which they have existing social connections, in which case the elements of $\beta_{2k(i)}$ that measure the effect of recent peer matriculations should be positive. Alternatively, strong students may dislike programs with academically weaker peer groups, in which case the elements of $\beta_{2k(i)}$ that measure the effect of admitted students with high school GPAs below i's GPA should be negative. We allow the effects of peer history to vary by whether the peers are of the same gender.

Parameterizing ex ante utility. Choice utility at the application stage, which we call ex ante utility, is a function of both ex post utility and the probability of admission. In the estimation, we follow Beuermann et al. (2023) in using a first-order Taylor approximation. That is, we condition on the probability of admission P_{is} by allowing it to enter into ex ante utility both additively and multiplicatively with the determinants of ex post utility. This approach allows for both the additive case where students explicitly trade off their ex post preferences against their admission probabilities and the case where students "discount" features of a program to which they are

[14] As a robustness exercise, we also estimate a specification with fixed effects at the level of high school-university pairs, discussed in Section 4.1. However, this specification adds a large number of parameters and the estimates are noisy.

unlikely to be admitted. The ex ante utility, $\delta_{is} + \varepsilon_{is}$, is then given by:

$$\delta_{is} + \varepsilon_{is} = \beta_{1k(i)} X_{is} + \beta_{2k(i)} H_{is} + \gamma_{1k(i)}(1 - P_{is}) + \gamma_{2k(i)}(1 - P_{is}) X_{is} + \gamma_{3k(i)}(1 - P_{is}) H_{is} + \varepsilon_{is} \quad (5)$$

We include the probability of admission as $(1 - P_{is})$ rather than the raw P_{is}, so the parameters $\beta_{1k(i)}$ and $\beta_{2k(i)}$ measure students' ex post preferences over program characteristics if admission were guaranteed, analogous to Equation 4. These parameters are equal across Equations 4 and 5 to maintain persistent preferences across the application and matriculation stages.

The parameter $\gamma_{1k(i)}$ measures how strongly students value applying to programs with very high admission probabilities, conditional on program characteristics and peer histories. The parameters $\gamma_{2k(i)}$ and $\gamma_{3k(i)}$ measure how severely students discount program characteristics and peer histories of programs due to low admission probabilities. For example, if the elementwise signs of $\gamma_{2k(i)}$ are the same as $\beta_{1k(i)}$, then a low admission probability leads students to "discount" program characteristics X_{is} that they would otherwise value. In other words, they value those characteristics and histories more for programs with guaranteed admission.

Persistent preference shocks. Thus far, we have treated the idiosyncratic error term in the ex post stage η_{is} as independent from the error term in the ex ante stage ε_{is}. This is tantamount to assuming that a student who has a positive preference shock for a certain major or a certain university at the time of submitting applications loses that shock (and draws a new one) by the time she is making matriculation decisions. This is a strong assumption. For example, it rules out students having a persistent preference for studying engineering, conditional on the observables in X_{is} and H_{is}.

We now relax this assumption and allow for persistent preference shocks across the two stages by using grouped fixed effects in the spirit of Bonhomme et al. (2019). This is a two-step approach. In the first step, we classify students into $K = 12$ types based on rich observables using k-means clustering (Appendix C.4). Twelve types allow for substantial heterogeneity.[15] To classify students into types, we use rich student characteristics to capture both horizontal and vertical differentiation

[15] In sensitivity tests, we find similar average results with other numbers of clusters. Twelve clusters produce reasonable cluster sizes. Fewer clusters would result in some very large clusters (above 20 percent of the sample), and more clusters would result in some clusters with less than 1 percent of the sample.

across students. For example, we include students' history of elective courses (not limited to grade 12 courses), allowing us to capture students' tastes and aptitudes. Many of these characteristics do not enter into the admission probabilities calculated in Section 3.5. Appendix C.4 provides the details of the k-means clustering and describes the resulting clusters. For example, the resulting clusters separate STEM-inclined strong students from STEM-inclined weaker students.

In the second step, we estimate Equation 3 separately for each cluster. This allows the preference parameters and the error terms ε_{is} and η_{is} to have persistent group-specific components. Specifically, the inclusion of major fixed effects and university fixed effects in X_{is} is equivalent to defining

$$\varepsilon_{is} = \zeta_{k(i),\text{major}(s)} + \zeta_{k(i),\text{university}(s)} + \tilde{\varepsilon}_{is}$$

$$\eta_{is} = \zeta_{k(i),\text{major}(s)} + \zeta_{k(i),\text{university}(s)} + \tilde{\eta}_{is}$$

Because the estimation proceeds separately for each cluster, this allows students of type k to have common persistent preferences for majors and universities across the two stages. For computational tractability, we maintain the assumption that $\tilde{\varepsilon}_{is}$ is independent from $\tilde{\eta}_{is}$, meaning that individual students still experience independent preference shocks over time, but only conditional on type membership.

There are at least two other possible approaches to relaxing the assumption of independence of ε_{is} and η_{is}. First, one might include detailed student observables interacted with program fixed effects directly in X_{is}. This approach would allow a more granular structure for the persistence of the preference shocks. With 305 programs, however, this substantially increases computational costs. In other applications with similarly structured high-dimensional choice sets, it has also been shown to lead to poor out-of-sample fit than more parsimonious specifications (Dingel and Tintelnot 2021). By contrast, the group fixed effects approach we adopt enables a dimensionality reduction that improves computational tractability while still preserving flexibility by fairly granularly defined student types. A second alternative approach is to continue to classify students into types for dimensionality reduction, but to allow the types to be inferred during the main estimation step rather than pre-specified based on observables in a first step. The estimation step would then include an expectation-maximization (EM) algorithm to estimate a mixture distribution that infers to which unobserved type a student belongs. The benefit of this approach is that it allows types to be a

more flexible function of both observables and unobservables. However, in practice, the number of unobservable types that could be feasibly allowed is small, because a larger type space substantially increases the computational burden. In our context, both of these alternative approaches are infeasible because we must access the OUAC data via a restricted server with limited computational capacity.

3.4 Identification of Peer Effects

Identifying peer effects is known to be a difficult task (Manski 1993; Sacerdote 2011). It requires distinguishing causal peer influence from other confounders that create correlations in the decisions within a network. Reverse causality can bias estimates if, for example, two friends both encourage one another to apply to the same college program. Bias due to unobservables may also arise at the level of an entire cohort, such as when a high school begins to offer a new college prep course due to attaining a critical mass of students.

We leverage two main sources of variation to identify peer effects. First, we rely only on within-high school variation in the set of peers a student is exposed to, and therefore the set of peers from whom a student can learn information. We use only across-cohort peer exposure to mitigate endogeneity concerns about common shocks within a high school cohort.

Second, we exploit variation in a student's exposure to information from peers who live in the same small neighborhood but attend a *different high school*. The neighborhood granularity is similar to US Census block groups. Because this peer information source originates outside the focal student's own school, it cannot be mechanically generated by unobserved shocks specific to the focal student's high school or cohort.

To implement the neighbor-based strategy, we subset the estimation to neighborhoods that satisfy two conditions. First, we require that students from the neighborhood attend at least two distinct high schools. Second, we subset to neighborhoods small enough to have reasonable probabilities of within-neighborhood social interactions: those with no more than 15 students per cohort. For identification, we use the neighbor's exposure to upper-cohort schoolmates who are not the focal student's schoolmates.

3.5 Admission Probabilities: Measurement and Heterogeneity

Estimating the role of peer influence requires a measure of the admission probabilities P_{is}. Our rich application and admissions data (Section 2.1) and Ontario's unusually transparent admissions system (Section 1) allow us to construct an informative measure of admission probabilities, which we then treat as observable and conditionally independent within a student across programs.

We estimate models of admission to each university-major-year triplet using the observed academic performance and admission decisions from the OUAC data. Because our notion of a program is a major category from Table C.1 rather than a specific sub-program, we sometimes pool sub-programs into a single cell if a university has, for example, multiple biology sub-programs. We then use the high school courses, grades, and admission decisions of all students who apply to a given major category at a given university in a given year to estimate probit models of admission as a function of student observables. For each university-major-year triplet, we estimate a flexible model of admission probability as described in Appendix C.2.

The estimated admission probability models perform well, which is unsurprising given the small gap between the information observable to the econometrician and what universities use for admissions decisions. The model fits are summarized in Figure A.4. The histogram shows the distribution of the area under the ROC curve (AUC) across all university-major-year cells. For 93 percent of the application sample, the AUC is fair or better (AUC ≥ 0.7), and for 72 percent of the application sample, the AUC is good or better (AUC ≥ 0.8).[16]

Conditional independence of admission probabilities. Ontario's institutional details motivate the assumption that admission probabilities are independent within a student across programs, after conditioning on observable transcript information. We provide a formal test of this assumption in Appendix C.3. If the assumption is violated, then admission outcomes at other programs to which the student applied should be predictive of admission to a given program. We fail to find evidence that they are predictive of admission, conditional on student transcript information. We also fail to find evidence that the identity of a student's high school is predictive of admission, conditional on transcript information.

[16]The ROC curve plots a prediction model's true positive rate against its false positive rate across all classification thresholds. The area under this curve summarizes predictive accuracy in a single number: it equals the probability that the prediction model assigns a higher predicted probability to a randomly selected positive case than to a randomly selected negative case.

Observed choices are consistent with admission probabilities. Our estimator in Section 3 requires two assumptions about admission probabilities. The first is that we have reliable estimates of admission probabilities, as demonstrated above. The second is that students rationally respond to this information when constructing their application portfolios.

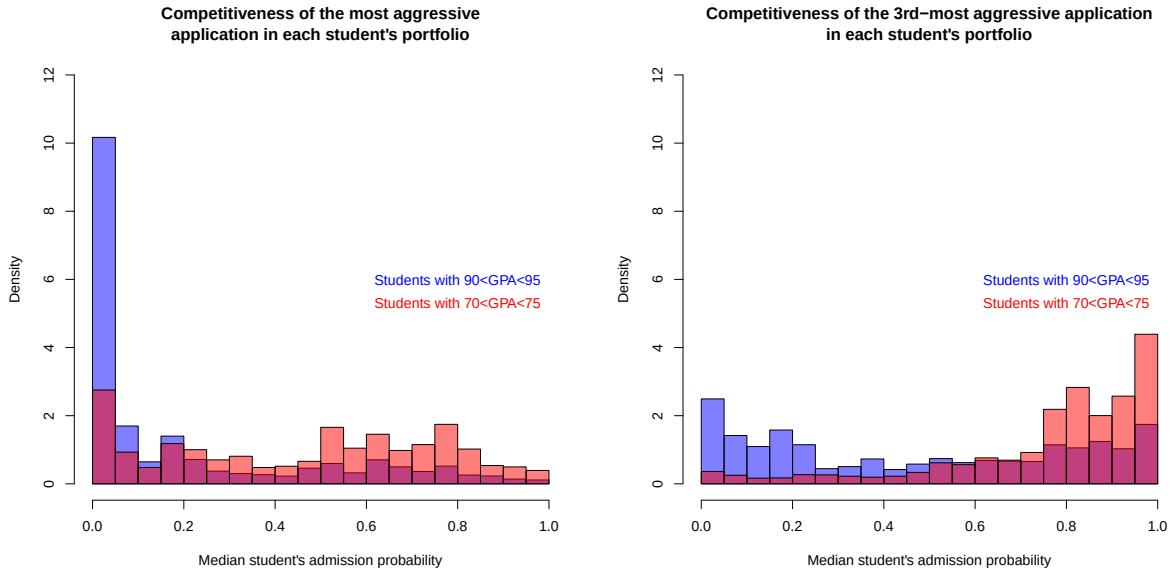
We now show that students' application behavior is consistent with a response to admissions probabilities. Using our admission probability estimates, we can compare application portfolios by students of various calibers. If students maximize total expected utility from an application portfolio as a function of program qualities and probabilities of admission as implied by rational behavior (Chade and Smith 2006), then stronger students will submit more aggressive application portfolios consisting of programs that have more competitive admissions. Indeed, Figure 2 shows that strong students with a high school GPA between 90 and 95 apply more aggressively than mediocre students with a high school GPA between 70 and 75.[17] In order to disentangle program competitiveness from the fact that stronger students have monotonically higher probabilities of being admitted to a program than weaker students, we define program competitiveness as the median predicted admission probability to a program across all students in that year, including students who did not apply to that program. Stronger students' applications are shifted to the left in Figure 2, indicating that they apply to programs for which the median student has a lower probability of admission.

The comparisons across student qualities in Figure 2 suggest that students rationally respond to their chances of admission. We also find no evidence of outright "mistakes." Students do not apply only to programs where they have a negligible chance of admission. The left panel of Figure 3 plots students' own admission probabilities at the programs to which they apply, binned by student GPA ventile. It shows that the vast majority of students—all but the bottom ventile—construct application portfolios that give them a higher than 50 percent chance of being admitted to at least one program. For most students, the chances of admission to at least one program are higher than 80 percent. The right panel shows a binned scatterplot of the probabilities of admission students could achieve if they applied to the 90th or 80th percentile "safest" programs, where "safe" is defined individually for each student. Comparing the left and right panels shows that students do not simply apply to the programs that are most likely to admit them; preferences over

[17] As explained in Appendix B.3, grades above 90 are considered an A+, and grades in the 70–79 range are considered a B– to B+.

programs are likely also at play. The right panel shows that students in the bottom ventile cannot substantially improve their chances of admission unless they apply to the extreme tail of high-admission-probability programs, which may well have utility below the outside option. Even the 90th percentile of possible admission probabilities for students in this ventile is below 40 percent. Taken together, Figures 2 and 3 suggest that students have rational responses to their chances of admission in constructing their application portfolios.

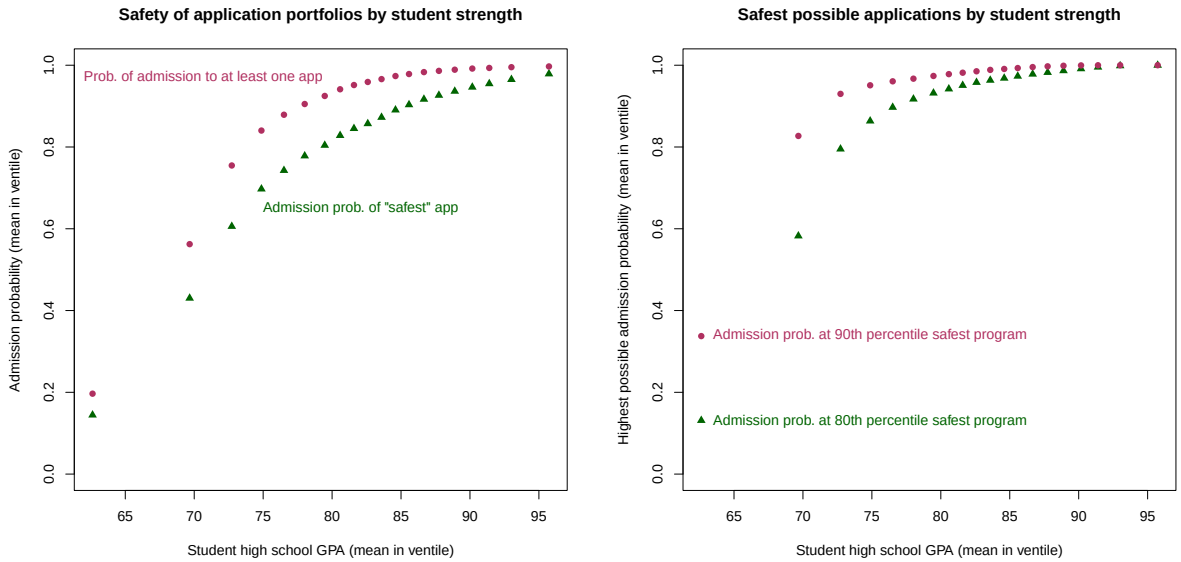
Figure 2: Aggressiveness of application portfolio by student quality



Note: Histograms of the competitiveness of the programs to which students apply. The horizontal axis measures the median student's admission probability to a program across all students in that year, whether or not they applied. Very strong students (blue) apply to more competitive programs with lower median admission probabilities than mediocre students (red).

Heterogeneity in application portfolios. There is heterogeneity across demographic groups in how aggressively students apply. Figure A.5 shows that students apply to more selective programs if they are from the Toronto area (as opposed to smaller cities or rural areas). This urban-rural gap is consistent with previously documented gaps in Ontario (Norrie and Zhao 2011) and geographic gaps in other markets (Fabre 2026; National Student Clearinghouse Research Center 2023). Similarly, men and students who attend large schools (where they are exposed to many peers) apply to more selective programs. For all three of these characteristics, the distribution of selectivity of applications is shifted to the left. The gap is not explained by differences in the composition of

Figure 3: Safety of application portfolio by student quality



Note: Binned scatterplots of student's own probability of admission by ventile of student quality. All students except those in the bottom ventile have application portfolios with a probability above 50 percent of admission to at least one program (left panel). Students do not necessarily apply to the programs that are most likely to admit them (left vs. right panel).

student quality, as the figure already conditions on a GPA in the top ventile. Figure A.6 shows that students in the top tercile of household income apply to more selective programs than students in the bottom tercile (top panel), though this pattern does not hold across the entire application portfolio for students in high quality ventiles (bottom panel).

There are also notable differences in applications to STEM programs by gender, consistent with Card and Payne (2021). The top panel of Figure A.7 shows that women are approximately 20 percentage points less likely to apply to STEM programs, even after conditioning on student quality (ventiles of high school GPA). After subsetting to students with a large number of STEM courses that qualifies them for admission to more STEM programs, the gap shrinks but does not close (top right panel).

4. Results

4.1 Choice Model Estimates

Table 2 shows our main estimates using both student decision stages. The table reports the average estimates across k-means clusters, weighted by cluster size. We include the admission probability as $(1 - P_{is})$ to interpret the coefficients on the raw variable as the ones corresponding to "safe" programs, as described in Section 3.3.

For example, the large and significant negative coefficient on distance suggests that when admission is guaranteed, students prefer programs closer to home. When admission probability is low (high $(1 - P_{is})$), as in the interaction coefficient column on the right, then students have an even stronger preference for nearby programs. Students from high-income zip codes, on the other hand, are less sensitive to distance.

Table 2: Maximum Likelihood Estimates - Neighborhood Interactions

	Raw variable		×(1–Pr(Admit))	
	Par	Std. Err.	Par	Std. Err.
(1-p)	2.219	0.031		
log(distance)	-0.500	0.005	-0.222	0.006
log(dist)*low income	-0.047	0.008	-0.003	0.009
log(dist)*high income	0.150	0.010	-0.002	0.011
Low income	0.001	0.059	-0.046	0.049
High income	-0.663	0.076	-0.046	0.060
student GPA	0.017	0.000	-0.003	0.000
Median Adm Prob	3.468	0.025	0.642	0.029
GPA * median adm prob	-0.046	0.000	-0.014	0.000
Female	0.187	0.051	0.404	0.042
Neighbor with Pr(interacted with peers who . . .)				
Applied in Cohort t-1	0.224	0.066	0.094	0.122
Applied in Cohort t-2	-0.099	0.213	-0.056	0.424
Admitted in Cohort t-1	-0.115	0.082	0.106	0.150
Admitted in Cohort t-2	0.184	0.281	0.074	0.527
Matriculated in Cohort t-1	0.066	0.123	0.035	0.250
Matriculated in Cohort t-2	-0.255	0.505	-0.049	1.575
Number of students	57,991			

Note: This table presents the parameter estimates from the maximum likelihood estimation using the neighborhood exposure probabilities. Columns (2) and (3) present the coefficients and standard errors on the raw variables listed in column (1). Columns (4) and (5) show the coefficient on the interactions between the respective column (1) variable and one minus the admission probability. Major and university fixed-effects are included in the estimation but omitted from the table.

We find that exposure to peers' applications meaningfully shifts students' own application portfolios. The coefficients on a neighbor's exposure to schoolmates who apply to a given program in

Table 3: Maximum Likelihood Estimates - Neighborhood Interactions

	Raw variable		×(1–Pr(Admit))	
	Par.	Std. Err.	Par.	Std. Err.
(1-p)	2.162	0.031		
log(distance)	-0.498	0.005	-0.224	0.006
log(dist)*low income	-0.049	0.008	0.000	0.009
log(dist)*high income	0.149	0.010	0.000	0.011
Low income	0.012	0.059	-0.059	0.049
High income	-0.658	0.076	-0.052	0.060
student GPA	0.017	0.000	-0.002	0.000
Median Adm Prob	3.459	0.025	0.820	0.029
GPA * median adm prob	-0.046	0.000	-0.017	0.000
Female	0.189	0.051	0.405	0.042
Neighbor with Pr(interacted with **same gender** peers who . . .)				
Applied in Cohort t-1	0.372	0.067	0.017	0.122
Applied in Cohort t-2	0.138	0.229	0.205	0.424
Admitted in Cohort t-1	-0.032	0.085	0.001	0.150
Admitted in Cohort t-2	-0.030	0.318	0.173	0.527
Matriculated in Cohort t-1	0.057	0.131	-0.003	0.250
Matriculated in Cohort t-2	-0.431	0.638	-0.012	1.575
Neighbor with Pr(interacted with **different gender** peers who . . .)				
Applied in Cohort t-1	0.022	0.087	0.118	0.157
Applied in Cohort t-2	-0.361	1.304	-0.278	2.105
Admitted in Cohort t-1	-0.174	0.116	0.132	0.192
Admitted in Cohort t-2	-0.088	3.602	0.114	5.220
Matriculated in Cohort t-1	-0.120	0.187	0.385	0.303
Matriculated in Cohort t-2	0.261	7.227	0.066	10.297
Number of students		57,991		

Note: This table presents the parameter estimates from the maximum likelihood estimation using the neighborhood exposure probabilities. Columns (2) and (3) present the coefficients and standard errors on the raw variables listed in column (1). Columns (4) and (5) show the coefficient on the interactions between the respective column (1) variable and one minus the admission probability. Major and university fixed-effects are included in the estimation but omitted from the table.

the preceding cohort is large and significant.[18] This is consistent with students interacting more with peers who are closer in age, conditional on being in the same classroom. The estimates in Table 2 use our primary definition of being in the same classroom: courses with enrollments of ten or fewer students. We find broadly similar results with the continuous measure of the probability that the students shared at least one physical classroom described in Section 2.3 (Table A.3).[19] With additional fixed effects at the level of high school-university pairs, we cannot reject that the estimates on the peer exposure variables are equal to the main specification (Table A.4).[20]

We further decompose the effect of peer applications into the effect of (a neighbor's) same-gender vs. opposite-gender peers to test for gender homophily (Hoxby 2000; Brenøe and Zölitz 2020; Dannemann 2020; Horrace et al. 2025). Table 3 shows that the baseline effect of peer applications is indeed driven entirely by same-gender peers.

The additional effect of peers being admitted is noisily estimated, and not statistically distinguishable from zero. This is consistent with admission probabilities already being observable to students. Peers who matriculate also have no statistically significant effect, suggesting that students do not have a preference for programs where their neighbor has a social connection. However, estimates using the continuous measure of students sharing a classroom do suggest a role for neighbors' social connections (Table A.3 coefficient on Matriculated in Cohort t-1). Taken at face value, this estimate suggests insider information diffused from a neighbor's older classmates to the focal student may play a role.

[18] In specifications including exposure to peers in up to four preceding cohorts, we find no effect of cohorts earlier than the two preceding cohorts. We therefore omit these from our main specification. In the data, the peer exposure to the third and fourth preceding cohorts are sparse.

[19] The point estimate on exposure to schoolmates who apply to a program in the preceding cohorts is 0.33 (standard deviation of regressor = 0.228), compared to 0.22 using the more conservative measure in our main specification (standard deviation of regressor = 0.102). The larger magnitude is partially driven by the main specification assigning some pairs who interacted to the "control" group. We therefore view the main specification estimates as a lower bound on the effect of peer exposure.

[20] Because the number of parameters increases rapidly with the number of schools, we use a small subsample of 30 schools that are large enough to have meaningful variation in peer exposures within school, subset to the largest six student clusters, and use only peer exposures in t-1. The top panel of table Table A.4 shows the estimates with these additional fixed effects. We cannot reject that the estimates on the peer exposure variables are equal to the main specification using this sample (bottom panel). However, the estimates with the additional fixed effects are substantially noisier, driven by the much larger number of parameters.

4.2 Policy Counterfactuals

We use the estimates from Table 3 to conduct counterfactual policy simulations. We evaluate the effect of changing students' exposure to information through peers on their application behavior. This requires solving for students' optimal portfolios under each counterfactual scenario.

Greedy algorithm for optimal portfolio choice. Finding the optimal portfolio is a combinatorial problem that typically requires evaluating millions of possible application portfolios.

To see this, suppose student i applies to a bundle of three programs, $B_i = \{1, 2, 3\}$. The associated admission probabilities are $\{P_{i1}, P_{i2}, P_{i3}\}$ and are assumed to be independent conditional on the student's observables. The ex post utilities of attending each program are $\{U_{i1}, U_{i2}, U_{i3}\}$, where $s \in \{1, 2, 3\}$ is numbered in the order of the student's ex post preferences: $U_{i1} > U_{i2} > U_{i3}$. The student's cost of applying to three programs is given by $c(3)$, where $c(\cdot)$ is increasing in the number of applications. The ex ante expected utility of this application bundle at the time of application is then:

$$R(B_i) = P_{i1}U_{i1} + (1 - P_{i1}) P_{i2}U_{i2} + (1 - P_{i1}) (1 - P_{i2}) P_{i3}U_{i3} - c(3) \tag{6}$$

This expression makes clear that the ex ante expected utility of a bundle of applications is not equal to the sum of the products of utility from attending each program multiplied by its admission probability. This is because a student only expects to realize the utility gains from at most one application, since she can enroll in at most one program even if admitted to multiple programs. The probability of admission to the second-most preferred program P_{i2} is only relevant if the student is not admitted to her first-most preferred program, which occurs with a probability of $1 - P_{i1}$.

A rational student does not simply apply to the programs with the highest ex post utilities, because applications to programs with zero (or sufficiently low) probability of admission are wasted. This presents a computational challenge. Staying with the example of a three-application bundle, to find the three optimal applications, the econometrician must calculate $P_{i1}, (1 - P_{i1}) P_{i2}$, and $(1 - P_{i1}) (1 - P_{i2}) P_{i3}$ for each permutation of three programs. With 21 universities and 18 major classifications, this results in 305 possible programs, and the number of possible (unranked) triplets exceeds 2 million.

This is an NP-hard problem, so in general it cannot be solved in polynomial time. However,

in our context, it satisfies conditions for a greedy algorithm that markedly simplifies solving for the optimal portfolio. The algorithm is derived by Chade and Smith (2006) under the crucial assumption that admission probabilities are independent across programs within a student. In such settings, the authors show that the optimal portfolio can be calculated in fewer steps as a function of programs' ex ante expected utilities $P_{is}U_{is}$. Their greedy algorithm requires evaluating only a subset of the possible permutations of programs by considering only one addition at a time to the previous step's application portfolio. With 305 possible programs, this reduces the problem to approximately 47,000 steps per student.

The algorithm finds the optimal set of applications in polynomial time, by proceeding as follows:

1. Start with the empty set of applications $B_{i,0} = \emptyset$ and set $n = 1$.

2. Among all not-yet-chosen possible programs, add any program s' such that $s' \in \arg\max_{s \in \mathcal{S} \setminus B_{i,n-1}} R(B_{i,n-1} \cup s')$, where $R(\cdot)$ is the overall expected payoff from the bundle of applications defined as in Equation 6 and $\mathcal{S}$ is the set of all programs.

3. If $R(B_{i,n-1} \cup s') - R(B_{i,n-1}) < c(n) - c(n-1)$, then stop. Otherwise, continue to step 4.

4. Update $B_{i,n} = B_{i,n-1} \cup s'$ and increment n by 1, and return to step 2.

A key characteristic of the optimal bundle is that it is more ambitious than the optimal single application. On the other hand, it is less aggressive than the optimal set of sequential applications, which consists of sequentially proceeding down a truthful ranking of the student's preferences. The algorithm also captures the notion of "reach" applications: it predicts that students will apply to some high-utility, low-admission probability programs if the rest of their portfolio is sufficiently certain (Appendix D provides an example).

The algorithm also provides the formal proof for the property that any program in the optimal application bundle has a greater ex post utility than any program with the same admission probability that is omitted from the application bundle (Chade and Smith 2006 Lemma 3). In other words, conditional on admission probabilities, a program in the application bundle is ex post preferred to a program not in the application bundle. This is the same basic insight we leverage for estimation. We use the greedy algorithm for finding optimal portfolios in the counterfactuals. To isolate the effect of peer exposure on the identities of programs in the application set, we fix the size of each student's application portfolio to the size observed in the data.

Gender gaps in applications. Our first set of counterfactuals concerns gender gaps in application behavior. In the data, women's application portfolios are less aggressive than men's (Figure A.5 middle panels). The gap is not fully explained by differences in admission probabilities: the gender gap is larger among students in the top deciles of high school GPA. The solid lines in the top panel of Figure 4 replicate this pattern with a simulation that uses the observed peer exposures in the data for men (red) and women (blue). Table A.5 columns 1–3 show the corresponding numeric values. In this baseline counterfactual simulation, in the 7th through 10th deciles of GPA, women's application portfolios consist of less selective programs on average than men's. In the top decile, the gap in average aggressiveness is 0.25 standard deviations. In the bottom half of the GPA distribution, women apply more aggressively than men, but the magnitude of the gap is small at 0.05 standard deviations, such that overall women apply less aggressively than men. The pattern is qualitatively similar using other measures of application aggressiveness, such as the selectivity of the most selective program in the portfolio.

The purple line with squares in the top panel of Figure 4 shows the results of a counterfactual analysis that shifts women's exposure to peers who apply to selective programs. Men's applications, the yellow line with circles, are unchanged because the counterfactual holds fixed their exposure to peers. Table A.5 columns 7–9 show the corresponding numeric values. We define selective programs as the 20 percent of programs with the lowest median admission probability.[21] Because the optimal portfolio is a function of both ex post utilities and admission probabilities, this counterfactual does not push women to apply "too aggressively" relative to their abilities (under the assumption that the reference group, men with similar abilities, is applying "correctly"). The counterfactual turns on women's exposure to peers applying to every selective program. We find that an intervention of this magnitude would close the gender gap in applications among the top decile of students, and would result in women applying more aggressively than men at every decile below that. We view this as a bounding exercise: endowing any student with exposure to *every* selective program is a large intervention, as the baseline average exposure is 0.01. Small interventions, such as endowing women with the same average exposure as men, result in negligible shifts in the gender gap of less than 0.001 (Table A.5 columns 4–6).

While our data do not allow us to follow students after they start college, we can measure

[21] The medians are calculated across all students to deal with selection into application, as in Section 3.5.

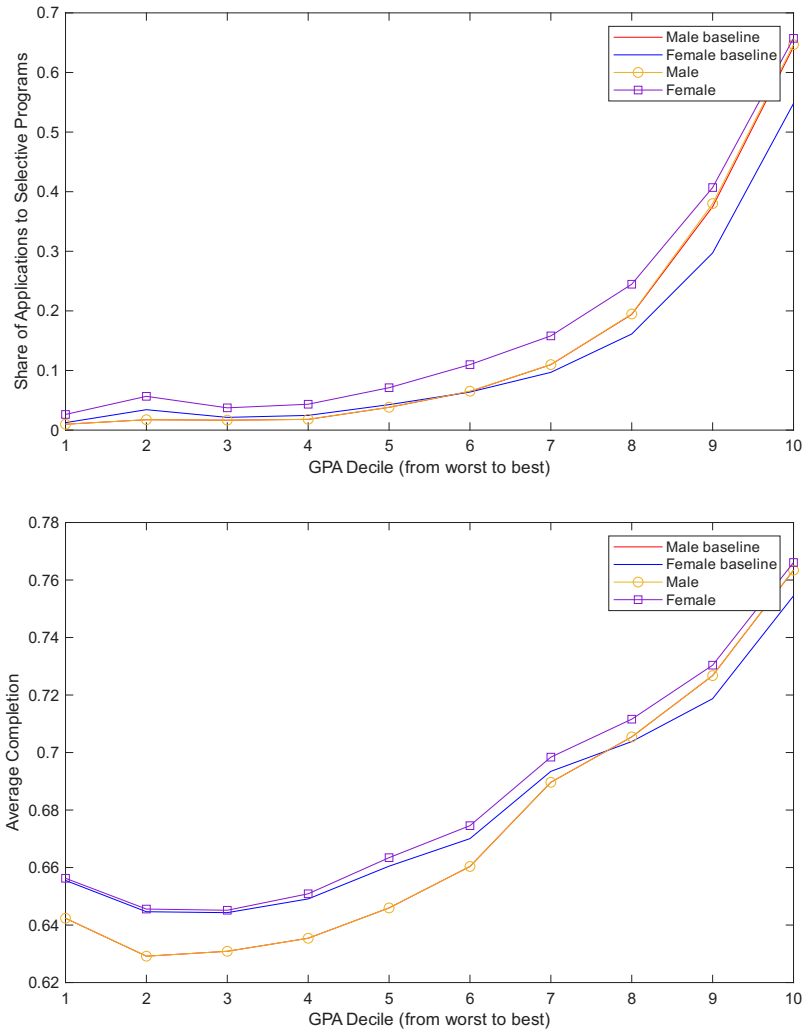
whether program selectivity is correlated with average program graduation rates, which are an important determinant of post-college outcomes. The bottom panel of Figure 4 reports the average program completion rates of programs students apply to in the baseline and counterfactual scenarios.[22] In the bottom seven deciles of GPA, women apply to programs with higher completion rates than men at baseline (blue line and yellow line with circles). In the top three deciles, women apply to programs with lower completion rates. Exposing women to peers applying to every selective program does little to shift the completion rates of women in the bottom half of the GPA distribution, but closes the gender gap in completion rates among the top third. Importantly, these findings should not be interpreted as a causal shift in individual students' college graduation probabilities as a result of peer exposure. Program completion rates are driven not only by causal effects of the programs themselves but also by student selection. Although the literature does not provide an estimate of the fraction of the variation in observed program completion rates that is driven by causal effects of colleges, a reasonable upper bound may be one third (Bound and Turner 2011). If causal effects of programs explain 20 percent of the variation, then our findings suggest that peer exposure could close 26 to 29 percent the gender gap in program completion rates among the top two deciles of students.

Women are also less likely to apply to STEM programs, even after conditioning on overall GPA and STEM-specific course performance (Figure A.7). Figure A.8 shows the results of endowing women with exposure to peers applying to every STEM program. Table A.6 shows the corresponding numeric values. Under the observed peer exposures, men's (solid red line) application portfolios are more heavily weighted toward STEM programs than women's (solid blue line). In the top GPA decile, the average fraction of STEM programs in men's portfolios is 70.9 percent, whereas for women is it only 59.5 percent. Endowing women with exposure to peers applying to every STEM program reverses the direction of this gap for all but the top deciles of GPA (purple line with squares). We find a qualitatively similar gap and reversal, though smaller in magnitude, when restricting the intervention to selective STEM programs (Table A.7 and Figure A.9).

Urban-rural gaps in applications. Students from small cities and rural areas apply less aggressively than students from large cities (Figure A.5 top panels). Table 4 columns 1–3 replicate

[22] We measure average program completion rates using data from the Ontario Ministry of Colleges and Universities (OMCU 2024). Completion rates are measured seven years after matriculation at the level of university-major category pairs.

Figure 4: Counterfactual Simulations: Gender Differences in Selective Programs



Note: Scatterplots showing the average share of applications in the portfolio that correspond to selective programs (programs in the top quintile of the distribution of median admission probability) of portfolio simulations by GPA decile (top graph) and the average program completion by GPA decile (bottom graph). The red and blue solid lines show the baseline simulation under the peer exposures observed in the data. The purple line with squares shows the counterfactual simulations endowing female students with exposures to peers for selective programs. Male counterfactual shares are equal to the baseline as we do not change their exposure to peers.

this pattern using a baseline simulation with the observed peer exposures. Students from outside large cities apply to less selective programs on average across their portfolio.[23] The most selective program in their portfolio is also less selective (Panel B), they apply to fewer selective programs (Panel C), and they apply to programs with lower completion rates (Panel D).

One driver of the urban-rural gap in applications to selective programs is that selective universities tend to be located in or near large cities, and students prefer nearby universities (Table 3). We therefore conduct a counterfactual that removes differences in students' distance to programs. Table 4 columns 7–9 show that this eliminates the gap. It also narrows the gap in observed program completion rates by approximately one third. Of course, this is an infeasible policy intervention that would require relocating students or universities en masse.

We therefore use this counterfactual only as a benchmark for an intervention that preserves locations but endows rural and small-city students with exposure to peers applying to every selective program. Columns 4–6 show that this would also close the geographic gap in applications to selective programs, with gap-closing effects similar to the relocation counterfactual. This illustrates the meaningful role of exposure to information through peers. While endowing students with exposure to every selective program is also a large intervention, it is a much lighter-touch intervention than relocation. For example, selective programs could send representatives or current students to talk to potential applicants in small-city and rural schools.

Program capacity constraints. In the counterfactuals simulations thus far, we hold students' admissions probabilities fixed at their observed values. The simulations allow students to reoptimize their application portfolios in response to changing peer exposures, but do not allow universities to adjust their admissions thresholds in response to changing numbers of applicants. We do this to isolate the role of peers, so that estimated changes reflect the counterfactual peer exposures rather than shifts in admission probabilities. However, if program capacity is costly or constrained, and students respond to universities tightening admissions thresholds, then our simulations will overstate the changes in student applications.

We therefore conduct a bounding exercise that takes program capacities as fixed, allowing admissions thresholds to adjust freely. This is computationally challenging and requires making

[23] "Large cities" is shorthand for the main metropolitan areas, which includes about 85 percent of Ontario's population. Students from these areas apply to an average of 2.3 selective programs, compared to 1.9 in small cities and rural areas.

Table 4: Counterfactual Simulations: Urban-Rural Differences

	Baseline	Exp. Rural to Select. Prog.	No Dist. Diff.
Panel A: Avg. Selectivity of the Portfolio			
Large Cities	0.707	0.707	0.743
Small Cities & Rural	0.734	0.702	0.745
Δ	-0.028	0.004	-0.001
Panel B: Avg. Select. of the Most Selective Program			
Large Cities	0.553	0.553	0.600
Small Cities & Rural	0.589	0.552	0,6012
Δ	-0.037	0.001	-0.001
Panel C: Avg N. of Selective Progr.			
Large Cities	0.462	0.462	0.442
Small Cities & Rural	0.421	0.577	0.408
Δ	0.042	-0.115	0.034
Panel D: Avg. Completion			
Large Cities	0.714	0.714	0.686
Small Cities & Rural	0.699	0.706	0.695
Δ	0.015	0.008	-0.009

Notes: Counterfactual simulations using optimal portfolio algorithm. The first column shows baseline simulations, where the optimal portfolio is constructed based on data. The second column changes peer connections for rural students and the third column eliminates geographic differences putting all programs at the average distance for all students. Selectivity is measured as the median student's admission probability; higher values correspond to less selective programs. Rural and small-city students apply to less selective programs than large-city students (column 1). Exposing rural and small-city students to peers who apply to selective programs closes or reverses the gap (column 2), an effect comparable to relocating students (column 3). Panel D shows average completion considering the programs in the chosen portfolio. At baseline, there is higher completion for students in large cities, that gaps closes approximately 40% when giving connections to selective programs to rural students.

additional rationality assumptions, as described in Appendix E.1. It also requires an allocation mechanism. We use a deferred acceptance algorithm assuming fixed program capacities equal to the number of admission offers observed in the data. There is no meaningful notion of application portfolios with a deferred acceptance algorithm, so we focus on final enrollments for this set of counterfactuals. To provide an apples-to-apples comparison to enrollments (rather than portfolios) with flexible capacity, we also simulate enrollments using a standard logit discrete choice setup in which students choose only among the programs where they meet the observed admissions thresholds.

The results suggest that the simulated effect of shifting peer exposures on demographic gaps under fixed capacities is not an over-estimate. When holding capacity fixed, we find that endowing women with exposure to peers who apply to STEM programs shrinks the gender gap in selectivity of STEM applications by 27 to 45 percent among the top two deciles of GPA (Table E.1). In the top decile, the gap shrinks by 27 percent, from 0.063 to 0.042. When allowing capacity to be fully flexible, the direction of the gender gap reverses. In counterfactuals simulating changes to the urban-rural gap in response to endowing rural students with exposure to peers who apply to selective programs, we find the gap narrows more under flexible admission thresholds than under flexible capacities for some deciles of GPA, but the reverse for others (Table E.2). Taken together, we view these results as merely suggestive, because simulations of enrollment do not map one-to-one onto our preferred counterfactual simulations using the full application portfolio. Nevertheless, they provide some reassurance that our conclusions about the effectiveness of peer exposure for closing demographic gaps do not require an assumption of flexible capacities.

5. Conclusion

This paper shows that high school students' college application and attendance decisions are affected by their peers. We leverage an institutional setting where admissions decisions are based almost exclusively on observables to recover students' preferences over college program characteristics, including their peers' histories with the program. Granular data allow us to identify off of students' indirect exposure to their neighbors' schoolmates.

We find that exposure to different peers can meaningfully shift college application behavior. The urban-rural gap in applications to selective programs can be equally effectively closed by

shifting rural students' peer exposures as by relocating students. The gender gap in applications to selective programs and STEM programs among high-ability students can also be narrowed by shifting women's peer exposures.

These results may help to explain the relationship between socioeconomic segregation and income mobility (Chetty et al. 2022). Exposure to peers who apply to colleges or majors with high earnings potential may be one mechanism through which students with low socioeconomic status benefit from socioeconomic diversity. Our results also suggest that information interventions about college may be more effective than implied by estimated treatment effects on the treated, because of the spillovers on other students exposed to the treated students.

References

Abdulkadiroğlu, A. and T. Sönmez (2003): "School choice: A mechanism design approach," *American economic review*, 93, 729–747.

Abramitzky, R., V. Lavy, and S. Pérez (2021): "The long-term spillover effects of changes in the return to schooling," *Journal of Public Economics*, 196, 104369.

Aguirre, J. and J. Matta (2021): "Walking in your footsteps: Sibling spillovers in higher education choices," *Economics of Education Review*, 80, 102062.

Allison, P. D. and N. A. Christakis (1994): "Logit Models for Sets of Ranked Items," *Sociological Methodology*, 24, 199.

Altmejd, A., A. Barrios-Fernández, M. Drlje, J. Goodman, M. Hurwitz, D. Kovac, C. Mulhern, C. Neilson, and J. Smith (2021): "O brother, where start thou? Sibling spillovers on college and major choice in four countries," *The Quarterly Journal of Economics*, 136, 1831–1886.

Anelli, M. and G. Peri (2019): "The effects of high school peers' gender on college major, college performance and income," *The Economic Journal*, 129, 553–602.

Avdeev, S., N. Ketel, H. Oosterbeek, and B. Van Der Klaauw (2024): "Spillovers in fields of study: Siblings, cousins, and neighbors," *Journal of Public Economics*, 238, 105193.

Barrios-Fernández, A. (2022): "Neighbors' Effects on University Enrollment," *American Economic Journal: Applied Economics*, 14, 30–60.

Beuermann, D. W., C. K. Jackson, L. Navarro-Sola, and F. Pardo (2023): "What is a Good School, and Can Parents Tell? Evidence on the Multidimensionality of School Output," *The Review of Economic Studies*, 90, 65–101.

Bobonis, G. J. and F. Finan (2009): "Neighborhood Peer Effects in Secondary School Enrollment Decisions," *Review of Economics and Statistics*, 91, 695–716.

Bonhomme, S., T. Lamadon, and E. Manresa (2019): "A Distributional Framework for

Matched Employer Employee Data," *Econometrica*, 87, 699–739, publisher: John Wiley & Sons, Ltd.

Bound, J. and S. Turner (2011): "Dropouts and diplomas: The divergence in collegiate outcomes," in *Handbook of the Economics of Education*, Elsevier, vol. 4, 573–613.

Brenøe, A. A. and U. Zölitz (2020): "Exposure to more female peers widens the gender gap in STEM participation," *Journal of Labor Economics*, 38, 1009–1054.

Breslow, N. (1974): "Covariance Analysis of Censored Survival Data," *Biometrics*, 30, 89.

Card, D. and A. A. Payne (2021): "High school choices and the gender gap in STEM," *Economic Inquiry*, 59, 9–28.

Chade, H. and L. Smith (2006): "Simultaneous search," *Econometrica*, 74, 1293–1307.

Chetty, R., M. O. Jackson, T. Kuchler, J. Stroebel, N. Hendren, R. B. Fluegge, S. Gong, F. Gonzalez, A. Grondin, M. Jacob, et al. (2022): "Social capital I: measurement and associations with economic mobility," *Nature*, 608, 108–121.

Dannemann, B. C. (2020): "Peer effects in secondary education: Evidence from the 2015 trends in mathematics and science study based on homophily," Tech. rep., Oldenburg Discussion Papers in Economics.

De Giorgi, G., M. Pellizzari, and S. Redaelli (2010): "Identification of Social Interactions through Partially Overlapping Peer Groups," *American Economic Journal: Applied Economics*, 2, 241–275.

Dingel, J. I. and F. Tintelnot (2021): "Spatial Economics for Granular Settings," *NBER working paper*.

Dynarski, S., C. J. Libassi, K. Michelmore, and S. Owen (2021): "Closing the gap: The effect of reducing complexity and uncertainty in college pricing on the choices of low-income students," *American Economic Review*, 111, 1721–1756.

Dynarski, S., A. Nurshatayeva, L. C. Page, and J. Scott-Clayton (2023): "Addressing nonfinancial barriers to college access and success: Evidence and policy implications," in *Handbook of the Economics of Education*, Elsevier, vol. 6, 319–403.

Fabre, A. (2026): "The Geography of Higher Education and Spatial Inequalities," Working Paper.

Fazekas, A. (2006): "Canadian Teaching and Cross-Border Training: An Interview With a Leading Education Consultant," *Science*.

Fu, C., J. Guo, A. J. Smith, and A. Sorensen (2022): "Students' heterogeneous preferences and the uneven spatial distribution of colleges," *Journal of Monetary Economics*, 129, 49–64.

Gale, D. and L. S. Shapley (1962): "College admissions and the stability of marriage," *The American mathematical monthly*, 69, 9–15.

Horrace, W. C., H. Jung, J. L. Presler, and A. E. Schwartz (2025): "What makes a classmate a peer? Examining which peers matter in NYC elementary schools: WC Horrace et al." *Journal of Population Economics*, 38, 52.

Hoxby, C. (2000): "Peer effects in the classroom: Learning from gender and race variation," Working Paper 7867, National Bureau of Economic Research.

Hoxby, C. and C. Avery (2013): "The missing "one-offs": The hidden supply of high-achieving, low-income students," *Brookings Papers on Economic Activity*, 2013, 1–65.

Hoxby, C. M. and S. Turner (2015): "What High-Achieving Low-Income Students Know About College," *American Economic Review*, 105, 514–517.

Lavecchia, A. M., H. Liu, and P. Oreopoulos (2016): "Behavioral economics of education: Progress and possibilities," in *Handbook of the Economics of Education*, Elsevier, vol. 5, 1–74.

Makles, A. (2012): "Stata tip 110: How to get the optimal k-means cluster solution," *The Stata Journal*, 12, 347–351.

Manski, C. F. (1993): "Identification of endogenous social effects: The reflection problem," *The review of economic studies*, 60, 531–542.

Mora, T. and P. Oreopoulos (2011): "Peer effects on high school aspirations: Evidence from a sample of close and not-so-close friends," *Economics of Education Review*, 30, 575–581.

National Student Clearinghouse Research Center (2023): "High School Benchmarks: National College Progression Rates for High School Graduates (2023 Report)," National Student Clearinghouse Research Center, Tableau dashboard.

Norrie, K. and H. Zhao (2011): "An Overview of PSE Accessibility in Ontario," Issue Paper 8, Higher Education Quality Council of Ontario, Toronto, ON.

OMCU (2024): "University Employment Outcomes, Graduation and Student Loan Default Rates," `https://www.iaccess.gov.on.ca/OsapRatesWeb/enterapp/home.xhtml`, ontario Ministry of Colleges and Universities. Data from the annual Ontario University Graduate Survey. Last modified March 26, 2024.

OUAC (2016): "Resource Guide. Ontario University Regional Dialogues 2016," .

Patnaik, A., M. Wiswall, and B. Zafar (2021): "College majors 1," *The Routledge handbook of the economics of education*, 415–457.

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

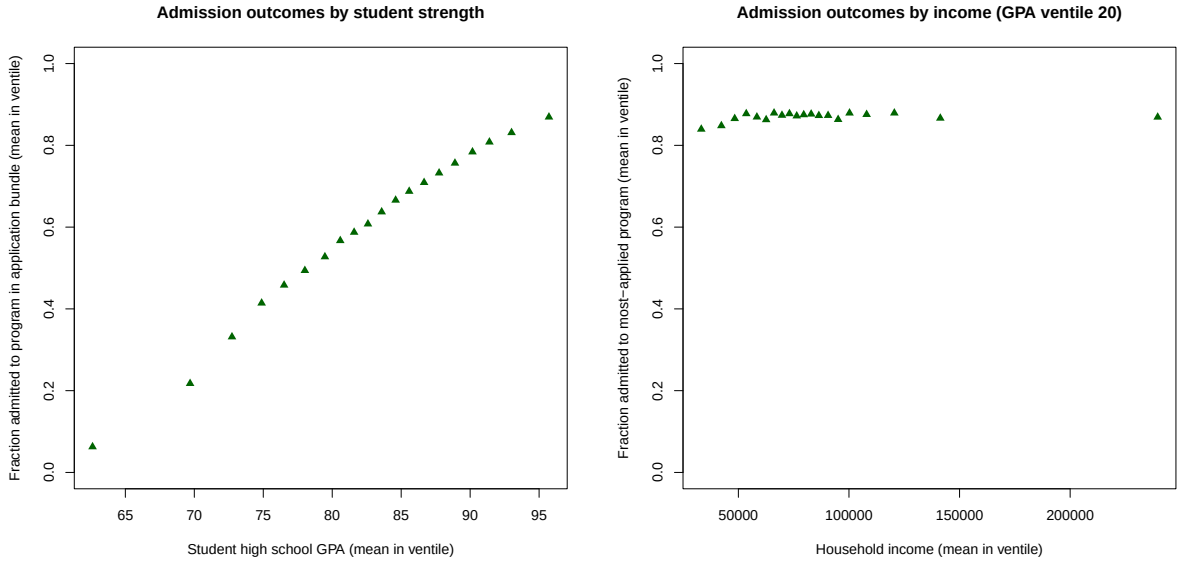
StatCan (2016): "Canadian students, tuition and additional compulsory fees, by level of study," Last Modified: 2023-09-06.

——— (2018): "Canadian and international tuition fees by level of study (current dollars)," Last Modified: 2023-09-06.

Zölitz, U. and J. Feld (2021): "The effect of peer gender on major choice in business school," *Management Science*, 67, 6963–6979.

A. Additional Figures and Tables

Figure A.1: Fraction of students admitted, by high school GPA and income



Note: Binned scatterplots of probability of admission by student characteristic. Probabilities are observed fractions of applicants admitted, not estimated probabilities. First panel bins students into ventiles of student quality (high school GPA). Second panel subsets to the top ventile of GPA and then bins students into ventiles of average household income in their zip code. There is a steep gradient in GPA, but a negligible gradient in income after conditioning on GPA.

Table A.1: Applicants per high school (means across schools)

Year	Applicants	Admitted	Women	Immigrants
1995	72	76%	54.5%	9.2%
2000	71	77.7%	56.5%	16.9%
2005	79	78%	56.2%	23.6%
2010	86	77.6%	54.8%	13.2%

Table A.2: Applications to each program (means across schools)

Year	Any apps	# of apps, if >0
1995	17.3%	2.9
2000	18.8%	3.6
2005	22.1%	4.1
2010	24.3%	5

Figure A.2: Full-time Undergraduate Tuition and Fees

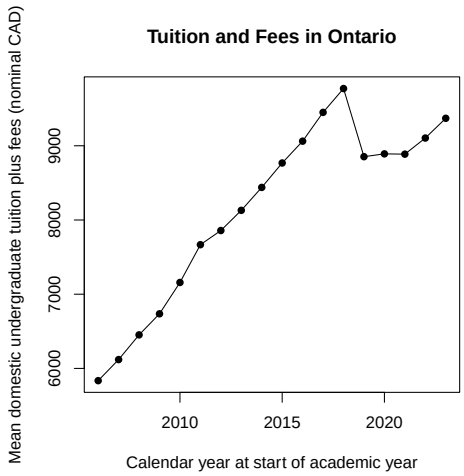
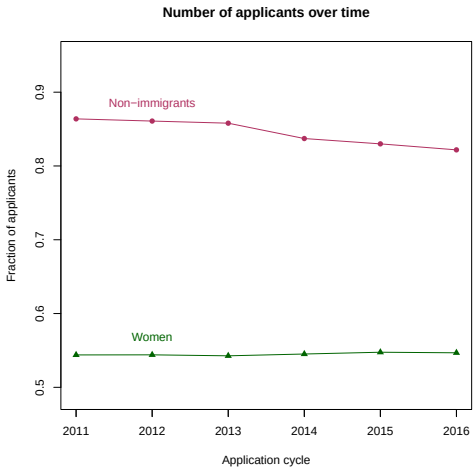


Figure plots mean tuition plus fees for Ontario undergraduates (domestic students only). Tuition is tightly regulated by the Ontario government and varies little across universities. The currency exchange rate is typically in the range of 1 CAD to 0.75–0.8 USD. Data from StatCan (2016).

Figure A.3: Applicant demographics over time



Fraction of applicants who are women (green triangles) and non-immigrants (red circles) in each year of the OUAC data.

Table A.3: Maximum Likelihood Estimates - Neighborhood Interactions using Continuous Measure

	Raw variable		×(1–Pr(Admit))	
	Par	Std. Err.	Par	Std. Err.
(1-p)	2.071	0.031		
log(distance)	-0.491	0.005	-0.219	0.006
log(dist)*low income	-0.049	0.008	-0.004	0.009
log(dist)*high income	0.149	0.010	-0.005	0.011
Low income	0.017	0.058	-0.039	0.049
High income	-0.655	0.075	-0.029	0.060
student GPA	0.017	0.000	-0.002	0.000
Median Adm Prob	3.530	0.025	0.684	0.029
GPA * median adm prob	-0.047	0.000	-0.015	0.000
Female	0.189	0.050	0.406	0.042
Neighbor with Pr(interacted with peers who . . .)				
Applied in Cohort t-1	0.332	0.043	0.046	0.071
Applied in Cohort t-2	0.154	0.126	0.009	0.234
Admitted in Cohort t-1	-0.062	0.046	0.089	0.077
Admitted in Cohort t-2	-0.064	0.157	0.104	0.281
Matriculated in Cohort t-1	0.123	0.057	0.047	0.106
Matriculated in Cohort t-2	0.107	0.232	-0.492	0.452
Number of students	57,991			

Note: This table presents the parameter estimates from the maximum likelihood estimation using the neighborhood exposure probabilities. Columns (2) and (3) present the coefficients and standard errors on the raw variables listed in column (1). Columns (4) and (5) show the coefficient on the interactions between the respective column (1) variable and one minus the admission probability. Major and university fixed-effects are included in the estimation but omitted from the table.

Table A.4: Maximum Likelihood Estimates - Neighborhood Interactions using Major and High School-University Fixed Effects

Panel A: Including High School x University FE

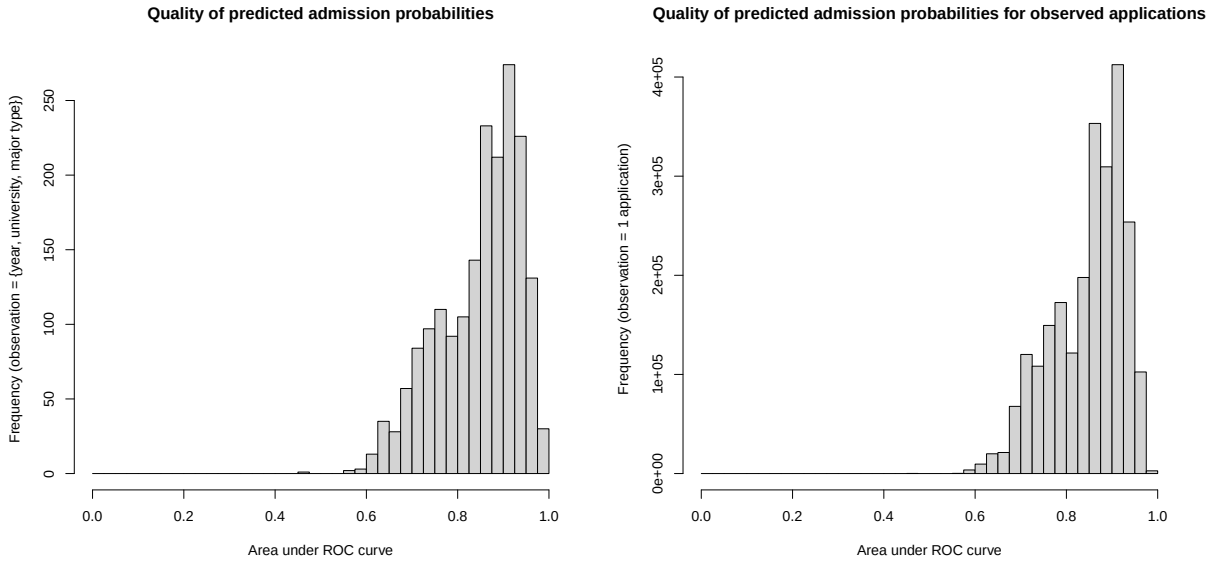
	Raw variable		×(1–Pr(Admit))	
	Par	Std. Err.	Par	Std. Err.
(1-p)	13.362	0.154		
log(distance)	-0.414	0.033	-0.516	0.033
log(dist)*low income	-0.043	0.057	0.173	0.058
log(dist)*high income	0.187	0.063	-0.006	0.064
Low income	0.108	0.464	-0.696	0.278
High income	-0.851	0.498	-0.015	0.304
student GPA	0.004	0.003	-0.128	0.002
Median Adm Prob	5.567	0.132	-1.992	0.150
GPA * median adm prob	-0.062	0.002	0.005	0.002
Female	0.469	0.384	0.360	0.229
Neighbor with Pr(interacted with peers who . . .)				
Applied in Cohort t-1	0.347	0.267	-0.221	0.440
Admitted in Cohort t-1	-0.304	0.381	0.046	0.579
Matriculated in Cohort t-1	-0.284	0.528	1.484	0.788
Number of students	40,598			

Panel B: Using Major and University Fixed-Effects

	Raw variable		×(1–Pr(Admit))	
	Par	Std. Err.	Par	Std. Err.
(1-p)	3.772	0.024		
log(distance)	-0.487	0.004	-0.230	0.005
log(dist)*low income	-0.057	0.007	0.008	0.007
log(dist)*high income	0.157	0.008	0.004	0.009
Low income	0.085	0.049	-0.087	0.037
High income	-0.680	0.062	-0.086	0.045
student GPA	0.022	0.000	-0.023	0.000
Median Adm Prob	4.248	0.019	0.498	0.022
GPA * median adm prob	-0.057	0.000	-0.009	0.000
Female	0.283	0.044	0.417	0.034
Neighbor with Pr(interacted with peers who . . .)				
Applied in Cohort t-1	0.250	0.054	0.084	0.095
Admitted in Cohort t-1	-0.142	0.068	0.109	0.116
Matriculated in Cohort t-1	0.050	0.102	0.076	0.179
Number of students	40,598			

Note: This table presents the parameter estimates from the maximum likelihood estimation using the neighborhood exposure probabilities, with a simpler specification including exposures of consecutive cohorts only. Panel A includes high school x university and major fixed effects. Panel B includes the same set of fixed effects used in the main results (university and major fixed effects). Columns (2) and (3) present the coefficients and standard errors on the raw variables listed in column (1). Columns (4) and (5) show the coefficient on the interactions between the respective column (1) variable and one minus the admission probability. Major and university fixed-effects are included in the estimation but omitted from the table.

Figure A.4: Performance of admission predictions



Note: Histogram of area under the ROC curve (AUC) of admission probability models. AUC is a standard measure of prediction model quality; an AUC of one indicates that the model perfectly predicts both positive and negative outcomes, while an AUC of 0.5 indicates that the model is no better than chance. The left panel has the university-major-year triplet as the unit of observation. The right panel has one observed application as the unit of observation, i.e. it is weighted by program popularity.

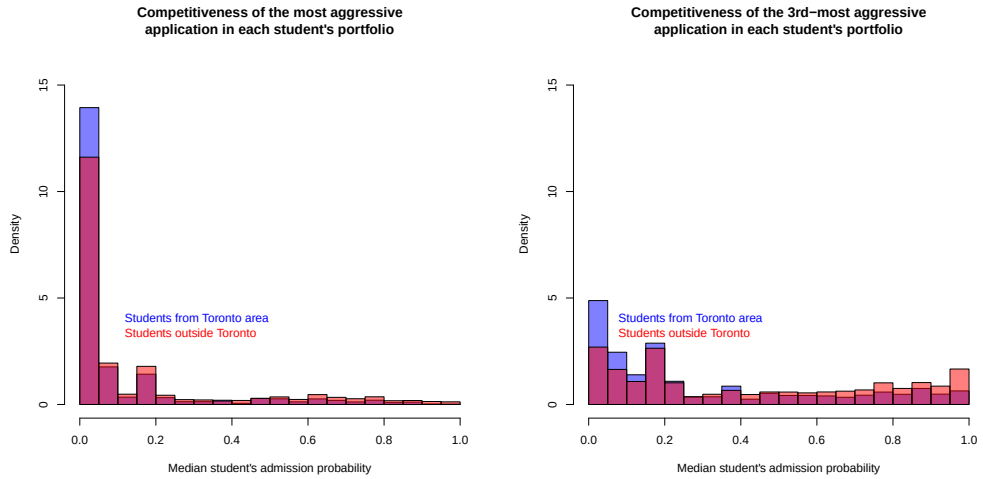
Table A.5: Counterfactual Simulations: Gender Differences

	Avg. Selectivity of the Portfolio								
	Baseline			Mens' Avg. Exp. to Women			Exp. to All Select. Progs. for Women		
GPA decile	Male	Female	Δ	Male	Female	Δ	Male	Female	Δ
1	0.874	0.859	0.015	0.874	0.859	0.015	0.874	0.850	0.024
2	0.877	0.85	0.027	0.877	0.850	0.027	0.877	0.837	0.040
3	0.873	0.866	0.007	0.873	0.866	0.007	0.873	0.857	0.016
4	0.864	0.852	0.012	0.864	0.852	0.012	0.864	0.840	0.024
5	0.839	0.825	0.014	0.839	0.825	0.014	0.839	0.807	0.033
6	0.793	0.792	0.001	0.793	0.793	0.001	0.793	0.763	0.030
7	0.716	0.723	-0.007	0.716	0.723	-0.007	0.716	0.685	0.031
8	0.612	0.628	-0.016	0.612	0.628	-0.016	0.612	0.579	0.033
9	0.471	0.517	-0.046	0.471	0.516	-0.045	0.471	0.457	0.014
10	0.288	0.342	-0.054	0.288	0.341	-0.053	0.288	0.289	-0.001

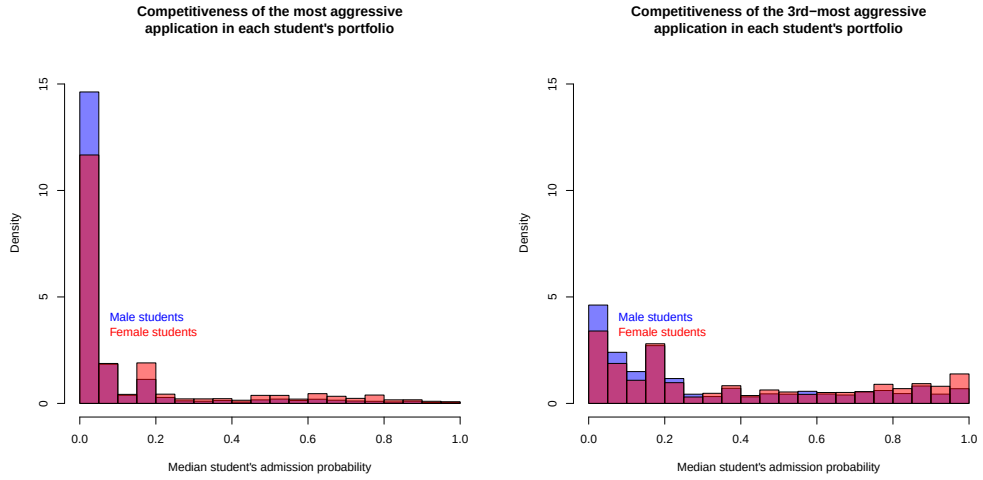
Notes: Counterfactual simulations using optimal portfolio algorithm. Selectivity is measured as the median student's admission probability; higher values correspond to less selective programs. High-ability women apply to less selective programs than high-ability men (column 1). Endowing women with men's average exposure to peers who apply to selective programs changes little (column 2). Exposing women differentially to peers who apply to selective programs closes the gap among high-ability students, and widens the reverse gap among low-ability students (column 3).

Figure A.5: Aggressiveness of application portfolio by student characteristics, among top ventile of students by GPA

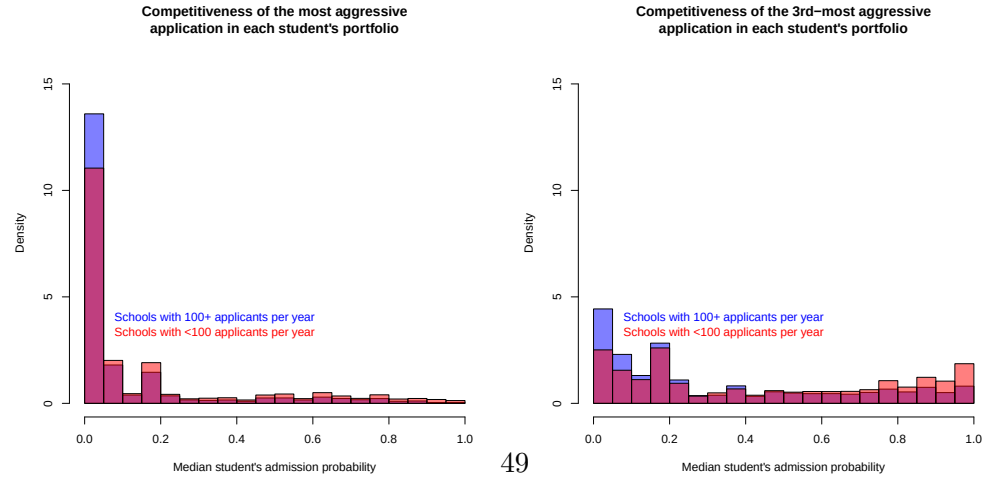
By geography



By gender



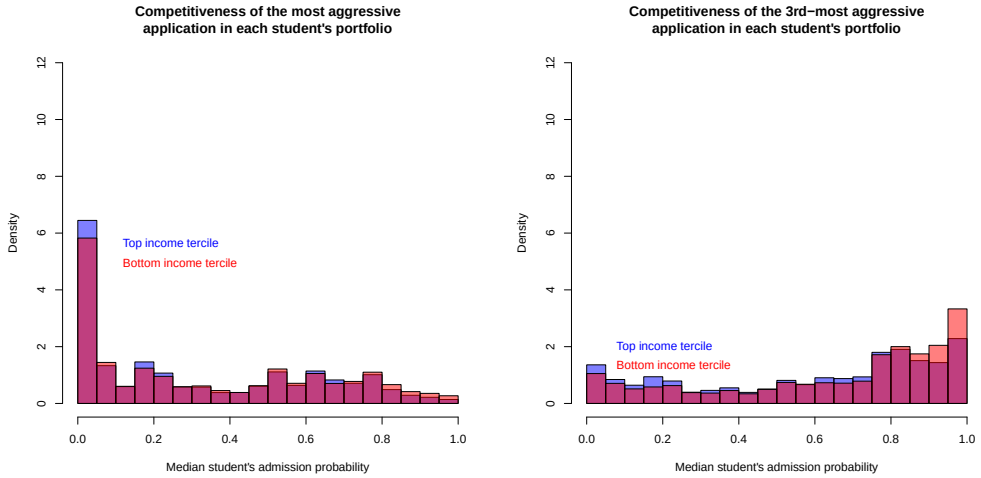
By school size



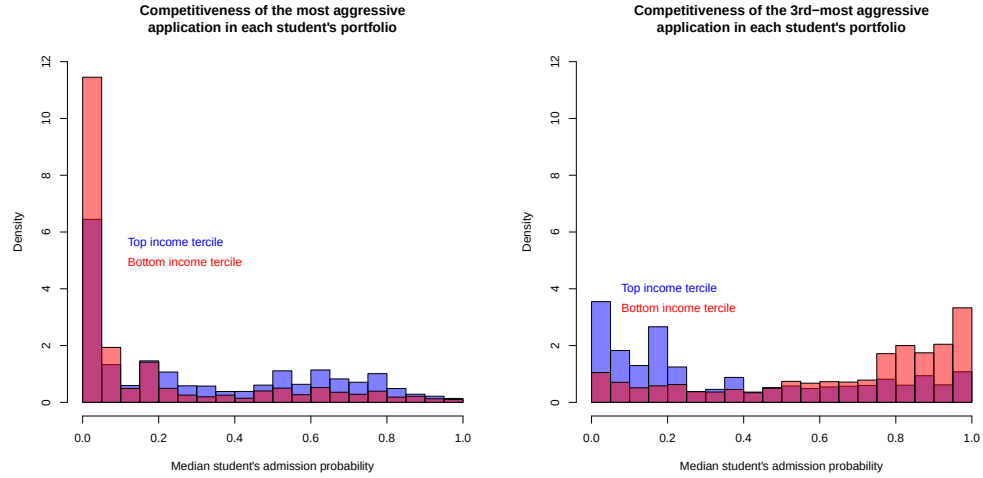
Note: Histograms of the competitiveness of the programs to which students in the top ventile of high school GPA apply. The horizontal axis measures the median student's admission probability to a program across all students in that year, whether or not they applied. Students from the Toronto area, students from large schools, and male students (blue) apply to more competitive programs with lower median admission probabilities than students from outside Toronto, students from small schools, and female students (red), even after conditioning on high school performance.

Figure A.6: Aggressiveness of application portfolio by household income

All students

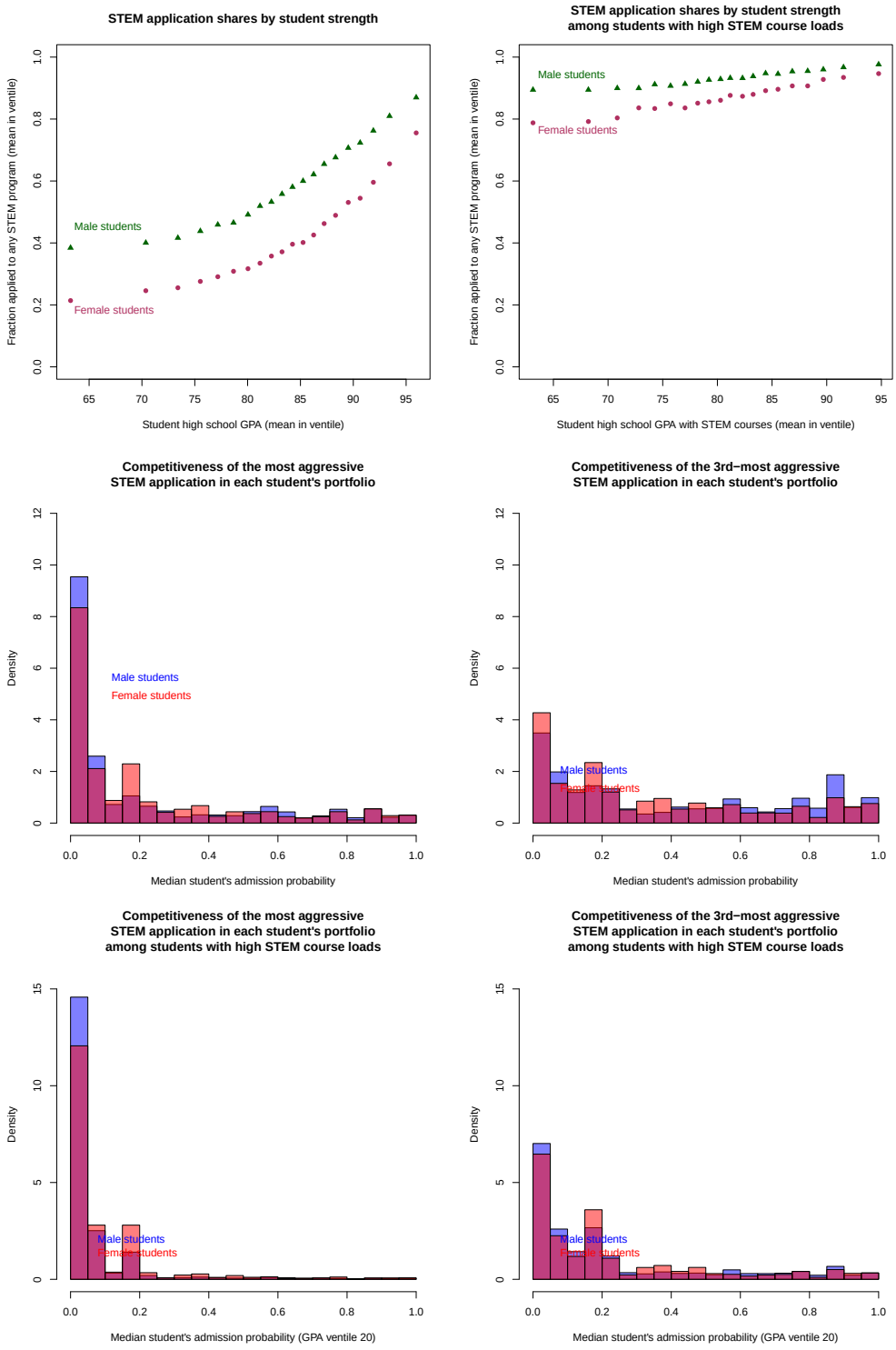


Top ventile GPA students



Note: Histograms of the competitiveness of the programs to which students in the top ventile of high school GPA apply. The horizontal axis measures the median student's admission probability to a program across all students in that year, whether or not they applied. Students who live in the top tercile of postal codes by household income (top panel, blue) apply to more competitive programs with lower median admission probabilities than students from the bottom terciles (top panel, red). After conditioning to the top ventile of high school performance (bottom panel), high-income students apply to less aggressive "reach" programs (bottom left panel) but the least aggressive applications in their portfolio remain more aggressive than those of low-income students (bottom right panel).

Figure A.7: STEM Applications by Gender



Note: Top panel: Binned scatterplots of the fraction of students in each GPA ventile who apply to any STEM program. Top left panel is unconditional on high school courses. Top right panel subsets to students who are well situated for STEM applications, having completed at least two science and at least two math courses in grade 12. Middle panel: Histograms of the competitiveness of the STEM programs to which students apply, conditional on any STEM application. The horizontal axis measures the median student's admission probability to a program across all students in that year, whether or not they applied. Bottom panel: Histograms of the competitiveness of the STEM programs to which students apply, subsetting two students who completed at least two science and at least two math courses in grade 12 and are in the top ventile of GPA. Male students (blue) are more likely than female students (red) to apply to any STEM program and to apply to a more competitive STEM program, even conditional on preparedness and top-ventile GPA.

Table A.6: Counterfactual Simulations: Gender Differences in STEM

	Share of STEM programs in the portfolio					
	Baseline			Exp. Women to STEM Progs.		
GPA decile	Male	Female	Δ	Male	Female	Δ
1	0.2619	0.1918	0.0701	0.2619	0.2939	-0.032
2	0.2425	0.1945	0.048	0.2425	0.31	-0.0675
3	0.2714	0.1854	0.086	0.2714	0.2958	-0.0244
4	0.3022	0.2327	0.0695	0.3022	0.3428	-0.0406
5	0.3403	0.2487	0.0916	0.3403	0.3561	-0.0158
6	0.383	0.2928	0.0902	0.383	0.4113	-0.0283
7	0.3569	0.2719	0.085	0.3569	0.3953	-0.0384
8	0.3659	0.2903	0.0756	0.3659	0.4124	-0.0465
9	0.4917	0.3722	0.1195	0.4917	0.4936	-0.0019
10	0.7086	0.5952	0.1134	0.7086	0.6962	0.0124

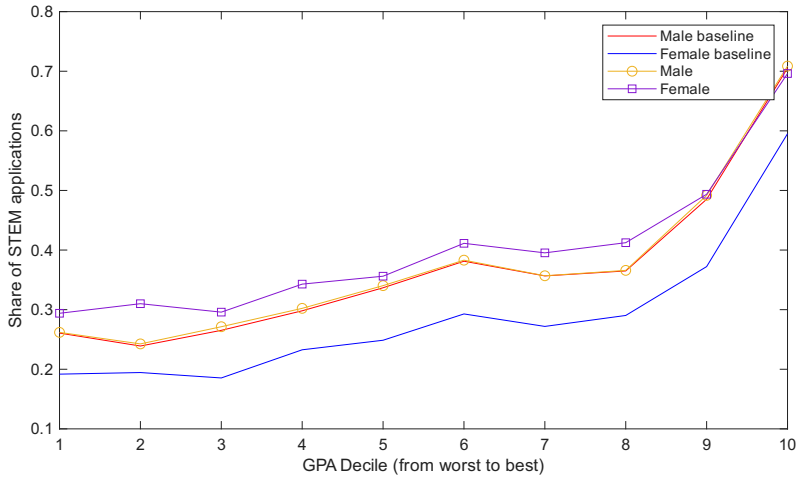
Notes: Counterfactual simulations using optimal portfolio algorithm. Women apply to STEM programs less than men (columns 1–3). Exposing women to peers who apply to STEM program reverses the gap (columns 4–6).

Table A.7: Counterfactual Simulations: Gender Differences in Selective STEM Programs

	Share of Selective STEM programs in the portfolio					
	Baseline			Exp. Women to Sel. STEM Progs.		
GPA decile	Male	Female	Δ	Male	Female	Δ
1	0.0099	0.0125	-0.0026	0.004	0.005	-0.001
2	0.0174	0.0343	-0.0169	0.004	0.008	-0.004
3	0.0166	0.0213	-0.0047	0.006	0.010	-0.004
4	0.0180	0.0247	-0.0067	0.011	0.027	-0.016
5	0.0382	0.0427	-0.0045	0.030	0.052	-0.022
6	0.0649	0.0638	0.0010	0.053	0.085	-0.032
7	0.1098	0.0969	0.0129	0.087	0.116	-0.029
8	0.1941	0.1612	0.0329	0.158	0.181	-0.023
9	0.3748	0.2972	0.0776	0.322	0.320	0.002
10	0.6430	0.5480	0.0950	0.588	0.588	0.001

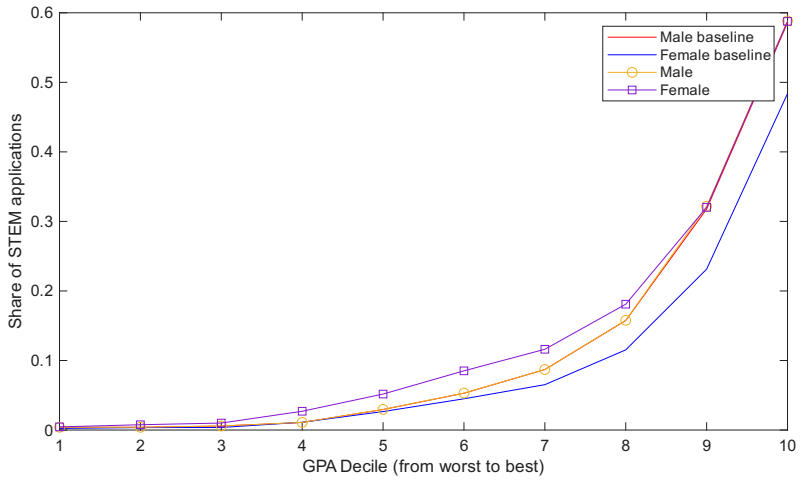
Notes: Counterfactual simulations using optimal portfolio algorithm. High-ability women apply to selective STEM programs less than high-ability men (columns 1–3). Exposing women to peers who apply to selective STEM program closes or reverses the gap (columns 4–6).

Figure A.8: Counterfactual Simulations: Gender Differences in STEM



Note: Scatterplot showing the average share of applications in the portfolio that correspond to STEM programs by GPA decile from portfolio simulations using the greedy algorithm described in Section 4.2. The red and blue solid lines show the baseline simulation under the peer exposures observed in the data. The purple line with squares shows the counterfactual simulations endowing female students with exposures to peers for STEM programs. Male counterfactual shares are equal to the baseline as we do not change their exposure to peers.

Figure A.9: Counterfactual Simulations: Gender Differences in Selective STEM



Note: Scatterplot showing the average share of applications in the portfolio that correspond to selective STEM programs (selective programs are in the top 20 percent based on their median admission probability) by GPA decile from portfolio simulations using the greedy algorithm described in Section 4.2. The red and blue solid lines show the baseline simulation under the peer exposures observed in the data. The purple line with squares shows the counterfactual simulations endowing female students with exposures to peers for selective STEM programs. Male counterfactual shares are equal to the baseline as we do not change their exposure to peers.

B. Higher Education in Ontario: Additional Context

This section describes Ontario's higher education environment, especially college admissions, during our study period of 2011–2016.

B.1 OUAC's Role in College Applications

The Ontario Universities' Application Centre (OUAC) is the central clearinghouse for all applications to Ontario universities' undergraduate programs. Students use OUAC's online platform to select undergraduate programs to which they wish to apply. Universities' admissions decisions are centrally communicated through the OUAC platform. Students accept or decline their admission offers through OUAC. Students also submit basic demographic information through this platform. Students are not responsible for uploading their transcripts; these are submitted to OUAC directly by high schools. High schools submit the student's academic record, consisting of grade 12 courses that the student has completed or is enrolled in, as well as any available final or interim grades in those courses.

OUAC also plays an important role in diffusing information about programs and their admissions requirements. Students can learn the admission requirements to each program by consulting eINFO, a publication that lists all available undergraduate programs in Ontario and their admission requirements for the upcoming application cycle. eINFO is published by OUAC and is distributed by high school guidance counselors and available online.[24] Figure 1 shows an example set of admission requirements, taken as a screenshot from one university's section of the 2015 eINFO booklet. Students observe the set of required and recommended courses for admission, reported as standard alphanumeric codes corresponding to courses in the Ontario high school curriculum. For example, ENG4U is grade 12 English, and MCV4U is grade 12 calculus. Also reported is the approximate minimum high school GPA of students admitted to the program in the previous application cycle. Students can estimate their probability of admission to a program based on whether they have completed the required and recommended courses, and by comparing their GPA to the approximate threshold from the previous cycle. GPA cutoffs typically remain stable within a program across adjacent years.

[24] Unlike in the US, guidance counselors play at most a small role in steering students' college applications, as described in Appendix B.4.

B.2 Tuition and Financial Aid

Tuition sticker prices are appreciably lower than in the US. Average sticker tuition and fees in 2015–2016, the last school year in our sample, were $8,768 in nominal Canadian dollars, or approximately $7,000 in nominal US dollars at a typical exchange rate (Figure A.2). With the exception of special professional programs, tuition is tightly regulated and varies little across universities or majors. Most students pay the sticker price unless they receive merit-based scholarships. Students can also apply for loans backed by the provincial government, but this is not done through OUAC's platform and is separate from admissions applications. There is no equivalent to the US centralized FAFSA financial aid form.

B.3 High School Curricula, Grading and GPAs

High school curricula are regulated by the Ontario government. In order to earn a high school diploma, a student must pass a set of compulsory courses and elective courses chosen from a preset menu. A full course load consists of eight courses per grade. The fraction of a student's course load coming from elective courses increases between grades 9 and 12. In grade 12, the only compulsory course is English. In grades 9–11, students must also complete a minimum number of courses in math, science, social science, French, and other electives.

Which grade 12 courses a student completes determines which undergraduate programs are likely to admit her. For example, engineering programs typically require calculus, one or two additional math courses, physics, and sometimes chemistry. Pre-med and life sciences programs typically require one or two math courses, biology, and chemistry. Humanities programs often do not require any particular grade 12 courses beyond the already-compulsory English course. We take as given students' high school courses, and study the downstream decision of college applications conditional on high school courses and grades. However, course enrollments provide crucial identifying variation in the set of peers a student is exposed to, even conditional on taking the same courses (Section 2.3).

Ontario high school grades are assigned on a 100-point scale. A grade of 90–100 corresponds to an A+; 80–89 is an A or A–; 70–79 is a B– to B+; 60–69 is a C– to C+; 50–59 is a D– to D+; and below 50 is a failing grade. Figure B.1 plots the distribution of high school GPAs in our data. More than half of students in the OUAC data—that is, those who apply to a four-year undergraduate

Figure B.1: High School GPA Among University Applicants

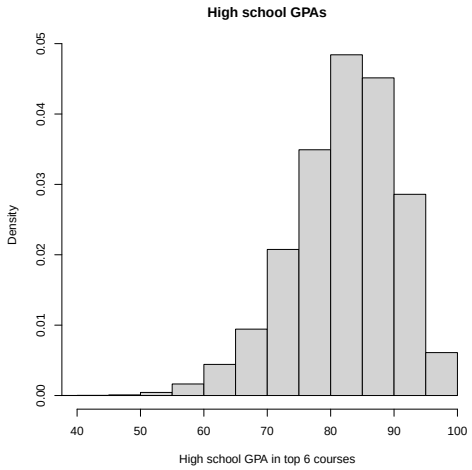


Figure plots the distribution of high school GPAs among students in our sample (those who submit at least one application to a four-year undergraduate degree in Ontario). GPA is calculated using the student's top six grade 12 courses, including English.

program—have a GPA above 80. Among the population of all high school students, this distribution would be shifted to the left. A student with a GPA in the 90s is considered excellent (approximately 15 percent of applicants). Fewer than half a dozen programs have minimum admission GPAs above 90. Even the most prestigious universities have minimum admission GPAs in the 80s for most majors.

B.4 Disparities Across High Schools

While resource disparities exist across schools and across districts, they are muted compared to the US. Funding levels are more homogeneous because all public school funding comes from a centralized province-level funding pool rather than from local property taxes. Moreover, there is less socioeconomic inequality in Canada, so a smaller share of students is severely disadvantaged. On the other hand, there is a high proportion of immigrants: more than a quarter of students have a first language other than English. The quality of teachers is likely less heterogeneous than in the US. The teaching profession is highly competitive in Ontario. In most cases, high school teachers must have an undergraduate degree plus an additional two-year degree known as "teacher's college". Teaching is viewed as a desirable and steady career, and admission to teacher's college is highly competitive (Fazekas 2006). Perhaps as a result of lesser heterogeneity in school resources, many

features of US K–12 education, such as charter schools, are absent in Ontario. Similarly, school guidance counselors play a much smaller role than in top US high schools, which is not surprising given that college admissions are based almost entirely on high school transcripts.

B.5 Neighborhoods and High Schools

Which high school a student attends is primarily, but not entirely, determined by the student's residence. This provides variation in the peer history a student's neighbor is exposed to in a different high school from that attended by the student, which we leverage for identification. Public schools in Ontario are run by school boards roughly corresponding to the geographic boundaries of municipalities. The public school boards' districts range in size from approximately 250 to approximately 250,000 K–12 students, with the largest being the Toronto District School Board. Students cannot cross school district boundaries. Even within a school district, the majority of students attend their neighborhood school.

Some high schools with special "magnet" programs draw some of their students from outside the neighborhood. The most common magnet program types are French immersion, where students take the majority of their courses in French; arts enrichment; science enrichment; and supplementary vocational training. These magnet programs typically enroll small cohorts of students within a larger school. Thus, even in magnet schools, the majority of students come from within the neighborhood and are enrolled in the regular curriculum. Admission to enrichment programs is typically competitive; admission to French immersion programs is not competitive but begins much earlier than high school, typically in first grade. Although they are in principle open to any student within the school district, magnet programs disproportionately enroll students from within the neighborhood or nearby neighborhoods. French immersion programs are found even in small school districts, because Canadian law guarantees most parents the right to choose between English- and French-language schooling for their children.

In addition to magnet programs, there is another mechanism through which students in the same neighborhood can attend different schools. Most municipalities have two parallel public school boards, a regular secular school board and a Catholic school board. Both are public and financed from provincial taxes. Two students living in the same apartment building can attend different schools even if neither of them attends a magnet program, if one attends a regular public

school and the other a Catholic school. In addition, 6 percent of students attend private schools (StatCan 2018).

C. Data Appendix

C.1 Classifying Majors of Study

Ontario universities offer a wide variety of undergraduate programs and majors, which we must aggregate for tractability. In 2010, for example, the universities offered over 900 distinct undergraduate programs, many of which allowed applicants to further specify a subfield of study. Some of the offered programs are quite general, such as mathematics. Others are highly specific. For example, Trent University offers separate programs for Indigenous Studies and for Indigenous Environmental Studies.

To make the analysis tractable, we classify undergraduate programs into 18 categories according to the field of study and admission requirements. Programs with close fields are grouped together, such as mathematics and statistics. Less similar fields are also grouped together if the admission criteria are similar and a sufficient number of universities do not clearly distinguish between the fields, such as computer science and engineering. In a few cases, a field that is similar to other fields is nonetheless carved out into its own category due to high enrollment, such as psychology. Finally, any field that leads directly to a recognized professional license in Ontario is carved out into its own category, such as nursing or education. The classification algorithm uses the subfield text descriptions present for approximately half of the observations in the OUAC data. This information is supplemented with text descriptions of each major code, collected from the eINFO booklets published by OUAC that students use to obtain information about undergraduate programs.[25]

Table C.1 lists the categories into which we have classified majors, and the fraction of total applications made to each category. The largest categories in terms of applications are business, commerce, and accounting at 19 percent of applications; engineering and computer science at 15 percent; biology, life sciences, and pre-med at 13 percent; social sciences at 9 percent; and humanities at 7 percent. Only a small fraction of applications are to general or undeclared majors: 5 percent are to an unspecified science major or an unspecified liberal arts major, and 1 percent

[25]There is a unique two- or three-letter code associated with each major within each university. The level of specificity of the codes varies across universities and over time.

are to majors that we could not classify even at the level of science vs. liberal arts.

Table C.1: Distribution of applications across major categories

Major category	Fraction of total applications
biology/life sciences/medical	13.5%
business/commerce/accounting	19.5%
chemistry	0.6%
criminology/forensics/law enforcement	3.8%
education	3.8%
engineering/computer science	14%
environment/ecology/agriculture	1.2%
fine arts/performing arts/design	4.9%
general humanities/social sciences	2.7%
general science/other science	1.7%
general/arts and science/undeclared	0.7%
humanities	7.8%
math/statistics	1.6%
nursing	3.5%
physics	0.9%
psychology	5.2%
social sciences	9.3%
sports/exercise/kinesiology/nutrition	5.3%

C.2 Estimating Admission Probabilities

For each university-major-year triplet, we estimate a model of the form:

$$\Pr(\text{admit}) = \Phi\left[spline(\text{GPA}) + \beta_1\,(\text{took req'd course}) + \beta_2\,(\text{grade in req'd course})\right]$$

where the required courses and the relevant courses for computing the GPA are defined as a function of the major. For each major category, a course is defined as required if the majority of universities require it for admission to a program in that major category. All programs require a standard English course. The most common courses required in other programs are calculus, other math courses, biology, physics, and other science courses.[26]

[26]The following major categories are coded as requiring a mathematics course: mathematics/statistics, engineering/computer science, physics, business/commerce/accounting, and general science/other science. The following are coded as requiring a science course: biology/life sciences/medical, chemistry, physics, general science/other science, environment/ecology/agriculture, nursing, psychology, and sports/exercise/kinesiology/nutrition. Additional details

The high school GPA used for each student mirros the GPA used by universities for admissions decisions. It is the mean grade across the student's best six courses, including the courses that are required for admission to a given program. High school GPA enters the model as a cubic spline, with knots defined as a sequence ranging from the minimum GPA for an admitted student to the maximum GPA for a rejected student, respectively rounded up and down to the nearest 5 points (on a 100-point GPA scale) and with a sequence step size of 5. These knot ranges are defined for each university-major-year cell to ensure that there is variation in the outcome variable in each between-knot subset of the GPA distribution.[27]

We require a minimum of 100 observations in each university-major-year cell to estimate the model. In cases with fewer than 100 applicants, we add data from the two adjacent years or the four adjacent years for the same university-major pair. If there are still fewer than 100 applicants, we instead add data from the two or four adjacent years for the same university and related majors (for example, adding general science to biology). Even so, 0.1 percent of applications are to university-major-year cells for which we do not have enough observations to fit a reliable admission probability model.

The estimated admission probability models perform well, which is unsurprising given the small gap between the information observable to the econometrician and what universities use for admissions decisions. The model fits are summarized in Figure A.4. The histogram shows the distribution of the area under the ROC curve (AUC) across all university-major-year cells. AUC is a standard measure of prediction model quality; an AUC of one indicates that the model perfectly predicts both positive and negative outcomes, while an AUC of 0.5 indicates that the model is no better than chance. For 93 percent of the application sample, the AUC is fair or better (AUC ≥ 0.7), and for 72 percent of the application sample, the AUC is good or better (AUC ≥ 0.8).

C.3 Conditional Independence of Admission Probabilities

Table C.2 provides a formal test of the assumption admission probabilities are independent within a student across programs, after conditioning on observable transcript information. If the assumption is violated, then admission outcomes at other programs to which the student applied should be predictive of admission to a given program. We therefore re-estimate our admission probability

on course classification are available from the authors upon request.

[27] If there is no variation, the probit model becomes unstable.

models adding granular information on the student's admission to or rejection from programs in every 10 percentage point category of selectivity. Selectivity is defined as the median admission probability across all students, even those who did not apply. We then test the hypothesis that these admission outcomes variables are jointly significant predictors of admission. We repeat the same test adding high school fixed effects. We fail to reject that the additional information on high school identity and other admissions outcomes are not predictive of admissions after conditioning on transcript information. The left tail of p-values is disproportionately among programs requiring additional application materials, such as fine arts, performing arts, and some business programs.

In order to implement these tests, we limit the sample to program-year pairs with a minimum of 500 applicants. We also estimate the models at the more granular level of individual university sub-programs (without aggregating to the university-major pair level) and using richer transcript information than in our main estimates. Specifically, we add indicators for specific courses completed by the student and grades received. We estimate linear probability models (LPMs) rather than probits for computational tractability with many additional regressors.[28] The resulting R-squareds close to one show that the imperfect predictions in Figure A.4 are due to the aggregation and parsimony of the probit models rather than due to underlying uncertainty.

Table C.2: Conditional independence of admission probabilities

Statistic	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
R-squared	0.992	0.037	1.000	1.000	1.000
Sig. of high school FEs	0.439	0.319	0.109	0.494	0.697
Sig. of other admissions outcomes	0.301	0.289	0.029	0.256	0.443

Note: Admission probability models estimated at the granular program-year level for cells with at least 500 applications. Estimated as linear probability models using detailed information from high school transcripts. Last two rows report p-values of tests for the joint significance of high school fixed effects, and of controls for admission and rejection outcomes at other programs applied to by the same student. The tests fail to reject that admission probabilities are conditionally independent of the identity of the high school or admission outcomes at other programs.

C.4 Clustering Students Using k-means

We use k-means clustering to categorize students into different types based on detailed information on courses and grades from the students course data. The objective is to capture unobserved prefer-

[28] LPMs are not appropriate for our main probability estimates because the application portfolio choice model requires valid predicted probabilities bounded between zero and one.

ences that create correlation between the errors terms in the application stage and the matriculation stage.

The clustering algorithm uses rich student characteristics to capture both horizontal and vertical differentiation across students. Many of these characteristics do not enter into the admission probabilities calculated in Section 3.5. To capture overall student strength, we use the student's GPA ranking within their high school; the standard deviation of grades across courses (to capture whether the student is uniformly strong or only strong in some subjects); and the count of advanced-track courses. To capture taste for and strength in STEM, we use the student's GPA in their best six courses including all math and science courses; the minimum grade across all math courses; the count of math courses; and indicators for each of biology and physics.[29] To capture other taste heterogeneity, we use the student's count of arts courses; and the count of social science courses. We standardize all non-binary variables to achieve a similar scale.

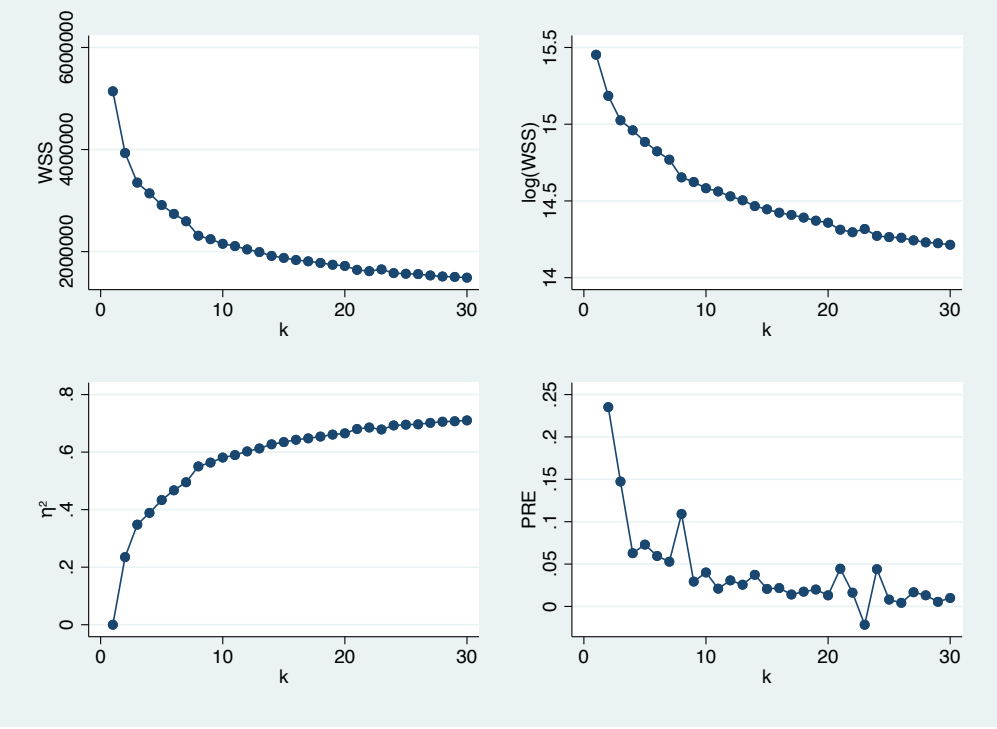
We use twelve clusters ($K = 12$). We use three commonly accepted measures to arrive at a minimum number of clusters. First, the within-cluster sum of squared errors (WSS), that measure how different the components of each cluster are with respect to the cluster mean. Second, the η^2 that measures the proportional reduction of the WSS for K clusters compared with the total sum of squares (TSS), and third, the proportional reduction of error (PRE) which illustrates the proportional reduction of the WSS for K clusters compared with the previous solution with $K-1$ clusters (Makles 2012). Figure C.1 shows the different measures for numbers of clusters ranging from two to thirty. The WSS graph does not show a clear kink, but overall results point to eight or nine clusters as a reasonable minimum. The η^2 criterion points to a reduction of the WSS by 60 percent and PRE to a reduction of about 12 percent from $K = 8$ to $K = 9$. We further expand to twelve clusters to ensure that there are no large cluster grouping together overly heterogeneous students.

Table C.3 shows descriptive characteristics for the twelve clusters, including the share of students in each cluster, the average GPA, and the counts of courses in relevant subject categories. Cluster 4 contains the highest-ability and most STEM-oriented students, and cluster 11 contains similarly STEM-oriented but lower-ability students. Clusters 1 and 10 contain high-ability but non-STEM–oriented students. Clusters 2 and 6 contains arts-oriented students. Clusters 3, 7, and 9 contain

[29] The Ontario high school curriculum only has one biology and one physics course available at the grade 12 level.

the lowest-ability students, of which Cluster 7 is the most STEM-oriented.

Figure C.1: k-means Clustering Statistics



Note: These plots show different statistics computed for each clustering procedure using k clusters as indicated in the x-axis. The top figures show the within-cluster sum of squared errors (WSS) and its log. The bottom graphs show the proportional reduction of the WSS compared with the total sum of squares on the left, and compared with the previous solution on the right.

Table C.3: Cluster Descriptives

Cluster	Fraction of Sample	GPA	U-Level Courses	Math Courses	Soc Sci Courses	Arts Courses	Physics Courses	Biology Courses
1	7.4%	85.502	6.808	1.613	2.401	0.187	0.095	0.246
2	5.0%	85.735	4.155	1.108	0.417	2.699	0.163	0.239
3	4.6%	68.981	3.046	0.025	0.575	0.575	0.026	0.171
4	13.5%	92.133	7.465	2.345	0.408	0.219	0.744	0.703
5	14.6%	87.752	5.302	1.885	0.342	0.270	0.263	0.379
6	5.5%	83.457	3.622	0.000	0.640	0.950	0.008	0.118
7	9.8%	71.254	4.970	1.572	0.412	0.245	0.269	0.411
8	11.9%	78.334	4.685	1.251	0.677	0.235	0.060	0.357
9	5.1%	73.923	4.995	0.073	2.233	0.188	0.010	0.108
10	3.9%	86.436	5.421	0.000	2.391	0.309	0.006	0.093
11	13.6%	80.892	6.570	2.226	0.274	0.128	0.616	0.593
12	5.2%	83.366	5.707	1.725	0.628	0.312	0.281	0.428

Note: Descriptives for students belonging to each of the twelve clusters used in the estimation. Column 2 shows the share of students belonging to that cluster. Column 3 reports the average GPA in the student's best six courses. The remaining columns report the mean number of courses in each relevant level or subject area. U-level courses are the most academically oriented among the standard course codes.

D. Additional Detail on Optimal Application Portfolio Algorithm

Section 4.2 describes the greedy algorithm proposed by Chade and Smith (2006) for finding the optimal application portfolio in settings like ours. The algorithm is capable of reproducing key stylized facts about college application behavior.

Most importantly, it predicts that optimal portfolios will frequently contain a "safe" program and a "reach" program. A safe program is one with a high probability of admission, but possibly a lower ex post utility of attending.[30] A reach program is one with a high ex post utility, but low probability of admission.

Consider a simple example with three possible programs whose ex post utilities and admission

Table D.1: Example of Chade and Smith (2006) Algorithm

Ex post utility rank	U_{is}	P_{is}	$P_{is}U_{is}$
1	10	0.01	0.1
2	8	0.999	7.992
3	5	1	5

[30] This is distinct from the definition of a "safety" program in Chade and Smith (2006), who define a safety program as one with high admission probability and (necessarily) low ex post utility. "Safety" programs are never optimal to include in the application bundle, but "safe" schools with only moderately lower ex post utilities can be optimal.

probabilities are as given in Table D.1. For ease of exposition, the programs are numbered in order of the student's ex post utility. The algorithm begins by taking the program with the highest $P_{is}U_{is}$, which is program 2 with $P_{i2}U_{i2} = 7.992$. At the next step, the algorithm checks the total expected payoff improvement from adding reach program 1 vs. safe program 3. The total expected payoff from adding program 1 is

$$R(\{1,2\}) = P_{i1}U_{i1} + (1 - P_{i1})\, P_{i2}U_{i2} = 0.1 + (1 - 0.01) \cdot 7.992 = 8.012$$

which is higher than the total expected payoff from adding program 3 of

$$R(\{2,3\}) = P_{i2}U_{i2} + (1 - P_{i2})\, P_{i3}U_{i3} = 7.992 + (1 - 1) \cdot 5 = 7.997$$

The optimal portfolio therefore consists of the safe and moderately preferred program 2 plus the reach program 1.

However, if program 2 becomes less safe because its admission probability drops to 0.9, then the optimal portfolio will include the safe program 3 instead of the reach program 1. The first step of the algorithm would still begin with program 2, with $P_{i2}U_{i2} = 7.2$. The total expected payoff from adding program 1 would become

$$R(\{1,2\}) = P_{i1}U_{i1} + (1 - P_{i1})\, P_{i2}U_{i2} = 0.1 + (1 - 0.01) \cdot 7.2 = 7.228$$

which is less than the total expected payoff from adding program 3 of

$$R(\{2,3\}) = P_{i2}U_{i2} + (1 - P_{i2})\, P_{i3}U_{i3} = 7.2 + (1 - 1) \cdot 5 = 7.7$$

E. Additional Results

E.1 Counterfactuals with Fixed Program Capacities

In the counterfactuals simulations in Section 4.2, we fixed students' admissions probabilities to those observed in the data. The simulations allow students to reoptimize their application portfolios in response to changing peer exposures, but do not allow universities to adjust their admissions thresholds in response to changing numbers of applicants, nor students to respond to these changes.

If program capacity is costly or constrained, then our simulations will overstate the changes in student applications.

We therefore conduct a bounding exercise that takes program capacities as fixed, which is equivalent to allowing admissions thresholds to adjust freely. To implement these counterfactuals, we introduce the additional assumption of no ex post regret. This assumption substantially reduces the computational burden because it implies that, after admissions decisions are realized, there is no program that a student did not apply to despite preferring it ex post to programs where she did apply, and where she would have been admitted. That is, students apply to (and ultimately enroll in) the program that gives them the highest ex post utility and to which they can be admitted. This simplifies the problem to a standard discrete choice over ex post utilities among a known feasible set, where each program has an admission probability of either one (in the feasible set) or zero.

We implement the fixed capacity exercise by first estimating a discrete choice model that includes only programs where the student's admission probability is above the observed threshold for admission that year. We then use a deferred acceptance algorithm to allocate students to programs (Gale and Shapley 1962; Abdulkadiroğlu and Sönmez 2003), determining the admission cutoffs endogenously for counterfactual simulations. The estimates are qualitatively similar to our main specification, including the effect of peer exposure. (Direct numerical comparisons are not straightforward because, unlike our preferred specification, the discrete choice model discards variables' interactions with admission probabilities.)

We do not use the known feasible set approach as our preferred specification for three reasons. First, it discards identifying variation from the application stage, whereas our preferred specification uses information from programs that the student applied to but did not attend. Second, it assumes away ex ante uncertainty about admissions thresholds, which is inconsistent with the observed fact of students applying to more than one program. Third, it does not allow solving for optimal application portfolios. However, this approach makes computations with fixed capacities feasible. Without it, counterfactual simulations would require iterating to convergence on the flexible admission probabilities. Because of the large number of programs, some with very few applications, counterfactual admissions thresholds are unstable and have poor convergence properties.

The deferred acceptance algorithm finds stable matches between students and programs by

admitting the top students to all programs and then letting them choose their most preferred program, and then iterating over the remaining students until all program slots are filled or there are no more students.[31] TablesE.1 and E.2 shows the results of the deferred acceptance algorithm using the observed admission thresholds (fixed admission probabilities) versus flexible admission thresholds with fixed capacities. There is only a negligible shift in students' program choices. This suggests that program capacities play only a small role in our main counterfactual simulations (Tables A.5 and 4).

Table E.1: Counterfactual Simulations with Fixed Capacities - Gender Differences

	Avg. Selectivity of Chosen Program					
	Constrained Capacities			Unconstrained Capacities		
GPA decile	Male	Female	Δ	Male	Female	Δ
1	0.976	0.981	-0.006	0.769	0.765	0.004
2	0.979	0.980	-0.001	0.756	0.754	0.002
3	0.945	0.967	-0.022	0.721	0.719	0.002
4	0.909	0.937	-0.027	0.701	0.701	0.001
5	0.741	0.710	0.032	0.689	0.687	0.003
6	0.701	0.700	0.002	0.674	0.671	0.004
7	0.685	0.680	0.005	0.648	0.649	-0.001
8	0.664	0.669	-0.005	0.631	0.629	0.002
9	0.568	0.577	-0.009	0.608	0.603	0.005
10	0.254	0.300	-0.046	0.573	0.566	0.007

Note: Simulated choices assuming fixed program capacities, using deferred acceptance algorithm to determine allocation. Students rank programs according to their utility and the algorithm determines admission. The table shows the average selectivity of the chosen program by GPA decile giving peer exposure to female students for STEM programs under the assumption of fixed versus flexible capacities. When assuming flexible capacities there is no need for an allocation mechanism, as each student just chooses their most preferred program.

[31] In our context, the supply of students can be exhausted before slots are filled because students can choose the outside option of not matriculating at any Ontario university.

Table E.2: Counterfactual Simulations with Fixed Capacities - Urban Rural Differences

	Avg. Selectivity of Chosen Program					
	Constrained Capacities			Unconstrained Capacities		
GPA decile	Urban	Rural	Δ	Urban	Rural	Δ
1	0.982	0.995	-0.013	0.792	0.787	0.005
2	0.979	0.987	-0.009	0.774	0.762	0.011
3	0.957	0.944	0.013	0.738	0.723	0.015
4	0.935	0.930	0.004	0.720	0.697	0.023
5	0.714	0.658	0.056	0.706	0.687	0.019
6	0.669	0.619	0.050	0.690	0.669	0.021
7	0.657	0.625	0.032	0.662	0.638	0.024
8	0.670	0.575	0.095	0.644	0.612	0.032
9	0.571	0.517	0.054	0.616	0.583	0.033
10	0.310	0.259	0.052	0.576	0.534	0.042

Note: Simulated choices assuming fixed program capacities, using deferred acceptance algorithm to determine allocation. Students rank programs according to their utility and the algorithm determines admission. The table shows the average selectivity of the chosen program by GPA decile giving peer exposure to rural students for selective programs under the assumption of fixed versus flexible capacities. When assuming flexible capacities there is no need for an allocation mechanism, as each student just chooses their most preferred program.