Peer Learning in College Applications

Ana M. Gazmuri
Toulouse School of Economics

Elena Prager
Simon Business School
University of Rochester
and NBER

December 2024

Abstract

Decisions about college are highly consequential, yet they are often made with poor information. This paper studies how information about college programs spreads through peer networks and affects application behavior using data from Ontario, Canada. We build and estimate a structural model of application and enrollment decisions. Ontario's unusual admissions rules and detailed microdata make all inputs into admissions decisions observable. This allows us to recast the NP-hard application portfolio problem as a tractable problem, by making a student's admissions probabilities to different programs independent conditional on observables. We also leverage student interactions across cohorts and with neighbors who attend other high schools to resolve standard challenges in estimating peer effects. We find that the probability of applying to a college-major pair increases significantly if a peer in an older cohort has applied to or enrolled in that same program, with stronger results if the peer enrolls. These findings shed light on the consequences of racial and socioeconomic segregation in schools and the long-term evaluation of information interventions.

1. Introduction

This paper studies how high school students' college application decisions causally respond to their peer networks. Peer effects are known to be an important determinant of behavior and outcomes, especially for teenagers. One of the most consequential decisions made by teenagers is whether and where to go to college, and what to study. Nevertheless, little is known about how exposure to one peer group or another affects decisions about college applications.

Do peers' application and attendance decisions cause students to apply to different colleges or majors? If so, then interventions that shift behavior among a small number of students may have large spillover effects. Perhaps more important, peer effects may play a role in information disparities affecting college applications and outcomes among less advantaged students (Cattan et al. 2023). This paper seeks to answer these questions and evaluate the contribution of peer effects to college application and attendance decisions.

Our approach relies on the estimation of a two-stage structural model in the context of Ontario, Canada. 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. Therefore, we model two utility functions, the ex ante utility that incorporates admission probabilities, and the ex post utility for the programs where admission has been granted. The key coefficients of interest are those on peers' history with a program. We are interested in the effects of whether the student has peers who applied to, were admitted to, or ultimately attended a program.

There are two key challenges to estimating this model. First, the absence of a centralized assignment system substantially complicates the identification and estimation of students' preferences over college options. In this context, students do not have incentives to truthfully list their most preferred programs. The observed applications are therefore a function of both true preferences and probabilities of admission. A student may not apply to her preferred program if she has little chance of being admitted, making it difficult to infer preferences from observed application choices. Furthermore, even if admission probabilities were known, solving for a student's optimal application portfolio constructed from among hundreds of possible programs is an intractable combinatorial problem.

We make the problem tractable using a theoretical result due to Chade and Smith (2006), who show that the optimal bundle can be calculated in fewer steps as a function of ex ante expected utility, defined as the probability of admission multiplied by the ex post utility of attending that program (if admitted). The Chade and Smith (2006) algorithm is not applicable to college application settings in general, but our context satisfies the required assumptions because admission probabilities are plausibly independent within a student across programs after conditioning on observables. Following Beuermann et al. (2023), we parameterize the ex ante utility as a more flexible function of admission probabilities and observables. The parameterization includes student observables, program characteristics, and program-specific peer group histories.

Accounting for the role of admission probabilities is typically infeasible. Admission probabilities depend on inputs that are (partially) observable to the student but not to the econometrician. For example, a student with an unobservably weak personal essay may apply to less selective colleges, which will bias downward the estimated preference for selectivity. Our unique empirical context and rich data make admission probabilities effectively observable, resolving this problem. Ontario's public universities are all public and tightly regulated, with limits on admissions practices. During our sample period, admissions are based almost exclusively on students' high school transcripts. Other factors commonly used in US college admissions are almost wholly absent in Ontario, including standardized tests, application essays, interviews, letters of recommendation, extracurriculars, and high school-specific admissions thresholds. Moreover, the admissions criteria are easily observable to students. This allows us to treat admission probabilities as observable to both the student and the econometrician, resolving the typical bias from application portfolios being a function of unobservable admission probabilities.

The second challenge is the standard identification problem when studying peer effects: two peers may make similar choices because the first peer influences the second or vice versa, because they both influence one another, or because they are exposed to a common unobserved shock. We mitigate the reverse causality concern by isolating the effects of older cohorts' choices on younger cohorts, so the direction of influence is clear. We are able to do this because of the structure of Ontario's high school curriculum leads to some mixed-cohort classrooms. In other words, students in their last year of high school could share a class with students one or two years below them. Additionally, we observe every class that each student has taken and when they took it, so we can identify student-specific interactions. We use the course date stamps to construct measures of which older-cohort students a student interacted with directly in a common classroom.

We use the history of application, admission, and matriculation outcomes among previous cohorts within their high school. This reduces concerns about reverse causality, but does not deal with endogeneity from common shocks. For example, two classmates may both be exposed to their high school hiring a new teacher who encourages students to apply to her alma mater. To mitigate concerns about common shocks, we also use information about neighborhoods and take advantage of students in the same neighborhood attending different high schools. We identify peer effects using the history of application, admission, and matriculation among previous cohorts in a high school attended by a student's *neighbor* rather than the student herself. We use a granular definition of neighborhood to capture neighbor pairs who are likely to know each other personally. Identification then requires that the focal student's choices are uncorrelated with the neighbor's upper-cohort schoolmates' decisions, except through information transmission via the neighbor.

Before presenting our structural results, we first establish the plausibility of peer effects in our setting using a descriptive analysis. We find that a student's probability of applying to a particular undergraduate program to which no one from her high school previously applied doubles after a student in the preceding cohort applies. Additionally, having someone from a preceding cohort choosing to attend the program is associated with further increases in the probability of application.

Our structural model estimates also show that peers play an important role. Interacting with a peer who applied to a given program significantly increases the probability of application, and this effect diminishes quickly with distance between cohorts. Whether the peer is admitted has little effect on the focal student's choice, consistent with our argument that admission probabilities are known to students in this context. On the other hand, if a student in the previous cohort chooses to matriculate, that further increases the probability of application.

These results have implications for spillover effects in policy design. Policy-makers looking to design effective interventions for directing students into high-return undergraduate programs rely on an understanding of the determinants of undergraduate college and major choice. All along the pathway from college consideration to matriculation, students face complicated choices and may lack sufficient information to navigate the process. The literature documents that young people, particularly those from lower-income and/or immigrant families, may lack good information about the process of preparing for, applying to, and selecting a college. Students may fail to identify and apply to programs and institutions that are a good match for them Avery et al. (2014) and many students may fail to apply to an appropriate number and range of institutions, even though it would be beneficial for them to do so (Smith 2014). There is a documented dearth of information regarding college opportunities, disproportionately affecting high-achieving low-income students, that impede students from optimally applying to and selecting colleges. Hoxby and Avery (2013) show that low-income students are less likely to apply to selective colleges compared to high-income students of similar achievement. They argue that low- and high-income students differ in their preferences, but also in their information set. This paper evaluates the effects of closing one contributor to the information gap: information from peers.

Peer interactions in the education process have been studied extensively along a number of dimensions. Most of this literature focuses on the effects of a student's peers on the student's achievement (Hoxby 2000; Sacerdote 2001; Imberman et al. 2012; Ellison and Swanson 2016). Peer effects on behavioral and social outcomes have also been studied. Gaviria and Raphael (2001) look at the influence of school peers in youth behavior such as drug use, alcohol drinking, and cigarette smoking. Mora and Oreopoulos (2011) study the influence of high school friends in educational aspirations and dropout intentions. School-level social interactions could also play an important role in decisions regarding college applications. Closer to this paper, Bobonis and Finan (2009) study peer effects on enrollment in additional education, and find a significant influence of neighbors' enrollment decisions under PROGRESA, a conditional cash transfer program in Mexico.

To our knowledge, there is little evidence regarding how peers influence college application decisions. Alvarado and Turley (2012) show that students with more college-oriented friends are more likely to apply to college. Furthermore, there is some evidence that students college applications follow the decisions of preceding cohorts from their own high school. This may be related to information diffusion, but also to issues of social belonging. Students may decide where to apply based on aspects of their own identity rather than academic success (Walton and Cohen 2007). Hoxby and Avery (2013) show that for high-achieving, low-income students who are geographically isolated from other high-achieving peers, college application choices mirror those of peers who are socioeconomically rather than academically similar. Altmejd et al. (2021) study the effects of siblings on college application decisions, showing that older siblings tend to influence choices of their younger siblings. In a paper closely related to ours, Abramitzky et al. (2021) show that high school students in Israel change their course selections and college matriculation decisions as a result of their peers learning new information about labor market outcomes. They use a compelling identification strategy that leverages a quasi-exogenous large shock to the peers' information set, in which their parents shift from being compensated at a uniform fixed wage to being compensated according to educational attainment. Our paper relies instead on smaller changes in information sets, on a scale more commonly experienced by high school students. We focus on diffusion of information that is a direct result of peers' actions, rather than information obtained by peers from a third party and then passed along. A closely related contemporaneous working paper by Borovickova et al. (2023) documents similar information diffusion in one school district in a reduced-form exercise. Relative to Borovickova et al. (2023), our paper formally models the students' full application portfolio decision and their choice among offers of admission.

The remainder of the paper is organized as follows. Section 2 describes the structure of high school curricula and college admissions in Ontario. Section 3 describes the data and shows motivating descriptive analyses. Section 4 presents the model of college application and choice, and describes how we parameterize it for estimation. Section 5 describes the estimates and describes the counterfactual policy evaluations we plan to conduct. Finally, Section 6 concludes.

2. College Applications in Ontario

2.1 Admissions Process

The majority of college-bound high school students in Ontario stay within Ontario for their post-secondary education. There are 21 public universities in Ontario during our sample period.[1] In Canada, 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. Ontario has no private universities. All universities are funded by the provincial government and subject to its rules governing admissions and tuition. For this paper, the most salient rules are ones that shrink the gaps between what is observable to universities, students, and the econometrician. First, Ontario prohibits universities from adjusting students' high

[1] This count includes only flagship campuses of multi-site universities.

school grades based on the location or perceived quality of the high school. This means that the high school performance we observe in the data is the relevant performance metric that universities use in their admissions decisions. Second, universities make admissions decisions based almost exclusively on students' high school transcripts, without the use of standardized test scores and negligible use of reference letters or extracurriclar activities. This means that there should be no unobserved student characteristics that are correlated with admissions probabilities. We describe these features in more detail in Section 2.2.

Like in the US, students apply to college during their final (twelfth) year of high school. The logistics of the application process are handled through a central clearinghouse used by all universities: the Ontario Universities' Application Centre (OUAC). Our data come from OUAC and are described in Section 3.1. Appendix B.1 describes the role of OUAC in the application process.

Unlike in the US, there is no equivalent to the centralized FAFSA financial aid form. Tuition sticker prices are appreciably lower than in the US. Average sticker tuition and fees in the last school year in our sample, 2015–2016 were $8,768 in nominal Canadian dollars, or approximately $7,000 in nominal US dollars at a typical exchange rate (Figure A.1). 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.

Universities typically admit students directly into a specific major. Admissions requirements vary by both university and major or program of study (see Section 3.3). For example, the same university may require a minimum grade in high school calculus for admission to its engineering major, but not require calculus at all for admission to its philosophy major. Students may apply to more than one major within the same university, in which case those applications are typically evaluated separately. The overwhelming majority of students (94% of applications in our data) declare a major at the time of application.[2]

There is a fee for submitting applications. The base fee ranges from $115 to $140 during our sample period, and allows a student to submit up to three application. Students can pay an

[2]Switching majors after matriculation is substantially rarer in than in the US. 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.

additional piece-rate fee of \$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.

2.2 Information Environment

Our setting is unusual in that universities' admission decisions depend almost exclusively on a student's academic performance in high school. Since we observe students' high school course selections and grades, the variables used to determine admission are observable to the econometrician. These observables and the observed admissions decisions allow us to construct reliable admission probabilities for any student applying to any program.

Each program specifies a set of required grade 12 courses, and uses grades in those courses and other grade 12 courses to determine admission. Students can learn the admission requirements to each program 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. 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.2 describes the grading scheme in more detail. 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 thresholds typically remain stable within a program across adjacent years.

The eINFO information booklets comprehensively list all program admission requirements. Admissions decisions are typically based entirely on the grade 12 GPA from the student's best six courses, including the courses required or recommended for admission to the program. Several factors that are important determinants of admission to US colleges are wholly absent from the admission process in Ontario. There is no standardized testing.[3] Universities are also prohibited

[3]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.

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 Guelph 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.

by provincial legislation from considering the identity or characteristics of a student's high school in the admissions decision.[4] 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.[5]

As a result, the econometrician observes nearly the entirety of the information used by admissions offices. The same is true for students. 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. There remains appreciable uncertainty about admission outcomes only in cases where a student's performance is close to the prior year's admission threshold, as the thresholds do fluctuate

[4]Disparities across high schools are also less pronounced than in the US, as described in Appendix B.3.

[5]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.

slightly year-to-year as a function of program capacity and application volumes. These features allow us to treat admissions probabilities as observable ex ante, even if the realization of the admission outcome is uncertain. Section 4.5 describes how we construct admission probabilities and shows that the constructed probabilities meaningfully predict student behavior.

2.3 High School Organization and Curriculum

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. 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 she could conceivably be admitted to. 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.

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. As a result, students in small schools are likely to interact personally with every student in their cohort and many students in adjacent cohorts. The majority of students, however, attend large schools where they may never be in the same class as some of their cohortmates. Half of students attend schools with cohorts of size 250 or larger. In large schools, two students taking the same course in the same semester are likely to be in separate course sections. In small schools, only one section of a given course may be offered per semester, particularly if it is an elective. This generates variation across students in their degree of overlap with observably similar classmates taking similar course loads. 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.

Beyond variation in exposure to peers within the same high school, our primary identification strategy leverages variation in exposure to the peer histories of neighbors attending other high schools. This is a feasible identification strategy because students living in the same small neighborhood can nevertheless attend different high schools. This can happen due to a combination of magnet programs and parallel tracks, described in detail in Appendix B.4. We leverage various sources of within-neighborhood variation high school attendance in constructing exposure to neighbors' peer histories in Section 3.4.

3. Data and Definitions

3.1 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. In an earlier sample beginning in 1998, we also observe applicant information and applications, but not admissions outcomes.

Ontario is Canada's largest province, with a population of approximately fourteen million, which is nearly 40% of Canada's total. We observe approximately 90,000 applicants per year for a total of approximately 546,000 applicants. Since we observe the universe of Ontario students who apply to Ontario universities, any sampling bias is negligible.

The OUAC data contain information on each key aspect of applications: the student's final year academic performance in each 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.[6] 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

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

not observe dates of the university's admission decision nor 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.

3.2 Sample Descriptives

There are approximately 90,000 applicants per year contained in our sample. Slightly more than half of applicants (54.5%) are women, consistent with women pursuing college degrees at higher rates than men. A substantial minority of applicants (15.5%) are foreign-born immigrants (Figure A.2), reflecting Ontario's large immigrant population; 29% of Ontario's population is foreign-born (Census 2011, Census 2016).[7] 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.

3.3 Majors of Study

The majority of students in Ontario (94%) declare a major at the time of application to undergraduate programs. 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 given category of major at a given university. Appendix C.1 describes how we aggregate programs into the eighteen 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% of applications; engineering and computer science at 15%; and biology, life sciences, and pre-med at 13%.

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

3.4 Peer Groups

Our identification strategy relies on interactions of students across cohorts that can lead to across-cohort information diffusion, and further diffusion through interacting with neighbors outside the high school. To study the effect of peer information on application decisions, we construct peer groups using information on how many high school classes, if any, a pair of students was enrolled in together. The high school curriculum in Ontario affords students a substantial degree of choice among courses. 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. Given that students in older cohorts apply to college before younger cohorts, this strategy mitigates reverse causality concerns, but we may still worry that the correlation could be explained by common unobservables within high school. To deal with common shocks within a high school, we take advantage students who live in the same neighborhood but attend different high schools. For each focal student, we use the student's *neighbors'* exposure to peers in the neighbor's high school that the focal student does not attend.

A full course load consists of eight courses per grade. The only required courses across the final two grades of high school are two English language courses. Students are also required to complete at least one math course and at least one science course, each of which can be chosen from a range of math and science courses offered by each high school.[8] For their remaining courses, students are permitted to choose from a wide range of humanities, social science, science, arts, and physical skills subjects as long as they choose at least one course from each of a handful of preset categories.

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 therefore share classrooms not only with classmates in the same cohort, but also classmates in the cohorts above or below theirs. We calculate the probability of each applicant having interacted with every student from her high school in the preceding three cohorts from the course enrollment data. The three-year cutoff corresponds to a student being in the same class during her freshman year of high school as a high school senior.

Probabilities of being in the same classroom can take on non-binary values because our measure

[8]For example, essentially the universe of high schools offer biology, chemistry, and physics as separate science courses, available separately at the grade 11 and grade 12 levels.

of a classroom identifier is noisy. Rather than observing 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 (i.e. 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 assign any pair of students who are enrolled in courses for which all these attributes match a positive probability of interaction.

For small inferred classrooms with enrollments of 25 or fewer, our baseline specifications assume that there is only one physical classroom sharing all these attributes. For inferred classrooms with enrollments above 25, we assume that there may be more than one physical classroom with matching attributes; 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:

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

In robustness checks, we use several other measures of the strength of student interactions. The first is simply the maximum probability of being in the same classroom across inferred classrooms for a given pair. In addition, we use the count of the number of inferred classrooms shared across two students, either using all shared inferred classrooms or subsetting to inferred classrooms with low maximum enrollments, such as 10 or 20 students.

3.5 Suggestive Evidence of Peer Effects

We begin by showing descriptive regressions consistent with student application behavior being responsive to information diffusion through peer networks. For now, these regressions treat a student's decision to apply to a given program as independent of her decision to apply to any other program. The structural model results in Section 5.2 relax this assumption by jointly estimating applications and program choices as described in Section 4.

The baseline regression estimates a student's probability of application to each university-major pair, as a function of the history of applications and outcomes by preceding cohorts in that student's high school. Our preferred specification uses histories for up to four cohorts back (three cohorts preceding the applicant), since the majority of high schools in the data span four grades.

For each high school, we compute the number of students in each cohort (if any) who have applied to a given college-major pair each year. Application histories are partitioned into discrete categories as a function of whether there were any applicants from a given high school to a given university-major pair in the preceding three or more cohorts. We use only the variation from cases where a high school's first application to the university-major pair occurred within the preceding cohorts. We use a look-back period of five years for this sample definition. Observations from high schools with a longer history of applications to the program are not used, because in such cases, there is no credible way to attribute positive cross-cohort correlations to new information shocks.

We estimate coefficients for each possible case: where the first application occurs exactly three cohorts back, two cohorts back, or occurs one cohort back. For these cases, we take all possible permutations of the preceding cohorts' binary application histories: at least one student applies three cohorts back, but no students apply two cohorts back nor one cohort back, denoted $A^{1,0,0}$; at least one student applies three cohorts back and again two cohorts back, denoted $A^{1,1,0}$; and analogously defined histories $A^{1,0,1}$ and $A^{1,1,1}$. The pattern of coefficients on these dummy variables captures the persistence and decay of information shocks, as discussed below.

For each applicant i potentially applying to university-major j from high school s in year t, we regress an indicator of application P_{isjt} on the vector of possible histories $\boldsymbol{A}$ and additional covariates. The regression takes the form:

$$P_{isjt} = \boldsymbol{\alpha} \cdot \boldsymbol{A} + \eta_s + \delta_{jt} + \varepsilon_{isjt} \tag{1}$$

where $\boldsymbol{\alpha}$ is the vector of coefficients on application history type dummies. By focusing on the cross-cohort relationships in the data, we eliminate concerns regarding simultaneity. To deal with cross-cohort correlations in unobservables at the school level, such as higher levels of applications from stronger high schools, we include high school fixed effects η_s. Finally, university-major-year fixed effects δ_{jt} control for fluctuations in popularity of specific program-university combinations.

Table 1 shows the results. There is a small positive effect of someone applying the year before, but the coefficients are much larger for the cases when people apply in several years. The magnitudes increase further with three years and four years in a row. These results are consistent with the idea that students transmit information to the peers with whom they interact during high school.

Table 1: Effects of the History of Previous Applicants

	Year of the first applicant			
	1 yr ago	2 yrs ago	3yrs ago	4 yrs ago
& nobody after	0.000400***	-0.001790***	-0.001359***	-0.000849***
	(0.000113)	(0.000119)	(0.000124)	(0.000126)
& every yr after		0.010363***	0.016175***	0.024309***
		(0.000164)	(0.000189)	(0.000180)
& yr -1			0.002209***	0.002328***
			(0.000210)	(0.000257)
& yr -2			0.000775***	0.001595***
			(0.000222)	(0.000251)
& yr -3				0.000960***
				(0.000244)
& yrs -1 and -3				0.005643***
				(0.000311)
& yrs -2 and -3				0.004041***
				(0.000293)
& yrs -1 and -2				0.006557***
				(0.000286)

Note: This table presents coefficient from a regression of the probability of application to a university-program pair on the history of previous applicants to that program from that high-school. All the coefficients come from the same regression, including program-year and high-school fixed effects.

4. 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 4.1 builds up to the empirical specification, starting with the theoretical underpinning of college application portfolios. Section 4.2 then describes the discrete choice problem faced by the student in deciding among offers of admission. Section 4.3 describes how we parameterize both stages in order to take the model to the data.

Finally, Section 4.4 describes the identification strategy. We leverage two main sources of variation to identify peer effects. First, we use within-high school variation in the set of peers a student interacts with, and therefore the set of peers from whom a student can learn information. We use only information transmission across cohorts of students to mitigate endogeneity concerns, such as those that might arise due to common shocks within a high school class. Second, we use student exposure to information from peers who live in the same neighborhood but attend a different high school.

4.1 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 applied and 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. 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. Therefore, underlying preference parameters are difficult to infer from observed application bundles. We make the estimation tractable using a simplification of a greedy algorithm for the portfolio choice problem along with a useful econometric approximation.

Greedy algorithm for portfolio choice. 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 ability and grades. 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:

$$R(B_i) = P_{i1}U_{i1} + (1 - P_{i1}) P_{i2}U_{i2} + (1 - P_{i1}) (1 - P_{i2}) P_{i3}U_{i3} - c(3) \tag{2}$$

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 π_{i2} is only relevant if the student is not admitted to her first-most preferred program, which occurs with a probability of $1 - \pi_{i1}$.

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 proposed greedy algorithm that markedly simplifies estimation. 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, compared to more than 2 million triplets of programs one must consider without the algorithm.

Ontario's unusual admissions environment is therefore key to making the college application portfolio problem tractable, because it allows us to treat the admission probabilities within a student as independent conditional on observables. In settings where unobservables such as reference letters are considered in admissions decisions, admissions decisions will be correlated within a student even after conditioning on observables. Such a correlation violates the assumptions required for the greedy algorithm, without which the problem is computationally intractable.

The algorithm finds the optimal set of applications 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 2 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.

This algorithm finds the optimal bundle of applications in polynomial time. The optimal number of applications is determined by the first application for which the bundle's overall expected payoff grows by less than $c(n) - c(n-1)$. 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 implies 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.

Simplifying the greedy algorithm for estimation. Chade and Smith's algorithm finds the optimal set of applications for a choice set of size N in at most $N(N+1)/2$ steps. With 305 possible programs, this can still require repeating nearly 47,000 steps per student for each calculation of the estimation objective function. We simplify the problem further for estimation using an insight from Beuermann et al. (2023). Their insight is to estimate a discrete choice model 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 ex post utility and admission probability, rather than ex post utility alone. This approach relies on the Chade and Smith (2006) algorithm's property that any program in the application bundle must have a greater ex post utility than a program with the same admission probability but not in the application bundle.[9]

Recasting the portfolio choice problem in terms of a maximization of a function of admission probability and true utility makes estimation tractable. We can now model a student's optimal application bundle as a discrete choice problem, where the choice utility is defined as a function of admission probability and ex post utility. Denote by $\delta_{is} + \varepsilon_{is}$ the student's choice utility of program s, where ε_{is} is an i.i.d. Type I extreme value idiosyncratic error term. In the empirics, we will parameterize this flexibly as described on page 24. 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 choice utility than every program 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.

Unfortunately, we do not observe the student's ranking among the programs in her application bundle. In the absence of an incentive-compatible central assignment mechanism, nothing in our

[9] Beuermann et al. (2023) leverage this property only for the special case of the ex post most preferred program in an application bundle. However, Chade and Smith (2006) Lemma 3 is more general, so we apply it to all programs to which we observe a student applying.

setting compels students to truthfully rank the programs to which they apply. We observe only whether a program was included in the bundle, not the student's truthful rankings within the bundle.

We therefore 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:

$$\Pr(B_i) = \prod_{s \in B_i} \frac{\exp(\delta_{is})}{\exp(\delta_{is}) + \sum_{s' \in \mathcal{S} \setminus B_i} \exp(\delta_{is'})} \tag{3}$$

This expression is a building block of the likelihood function we use for estimation.

The assumption of equal preferences is weaker than it first appears. This is because in our setting, making the assumption is approximately equivalent to estimation with all possible permutations of preference rankings within the bundle. It can be formally shown that Equation 3 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 (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, with slight variation across years as a result of small changes in universities' program offerings. The median student applies to just four programs (Section 2.1), and the 90th percentile student applies to seven programs.

Equation 3 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 (See Section 2.1), the size of observed application bundles is not very informative in practice, and we simply omit application costs from the estimation. 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.

4.2 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 4.1. The choice set is defined as the set of programs to which the student was admitted plus the option of not choosing any of those, so the relevant ex-post admission probabilities within the choice set are equal to one. This 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, so we pool both into a single outside option. For this reason, we allow the value of the outside option to differ by the quality of the student, measured by the average high school GPA.

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. 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

$$\Pr(t|A_i) = \frac{\exp(\nu_{it})}{\sum_{s \in A_i} \exp(\nu_{is}) + \exp(\nu_{i0})} \tag{4}$$

where ν_{i0} is the mean utility of the outside option. We allow the value of the outside option to be a function of student quality. This captures the fact that high-ability 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.

4.3 Parametrization and Estimation

We combine information from the observed application bundles (Equation 3) and observed matriculation decisions (Equation 4) to jointly estimate all utility parameters using maximum likelihood.

The total likelihood for student i is then:

$$\Pr(B_i)\Pr(t|A_i) = \left[\prod_{s\in B_i} \frac{\exp(\delta_{is})}{\exp(\delta_{is}) + \sum_{s'\in\mathcal{S}\setminus B_i}\exp(\delta_{is'})}\right]\left[\frac{\exp(\nu_{it})}{\sum_{s\in A_i}\exp(\nu_{is}) + \exp(\nu_{i0})}\right] \tag{5}$$

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.

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.

Parameterizing ex-post utility. 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 peers' history or 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 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:

$$\nu_{is} + \eta_{is} = \beta_1 X_{is} + \beta_2 H_{is} + \eta_{is} \tag{6}$$

The term β_1 measures a given type of student's valuation for a given set of program characteristics, in the absence of any peer history with the program. The term β_2 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 β_2 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 β_2 that measure the effect of admitted students with high school GPAs below i's GPA should be negative.

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

$$\delta_{is} + \varepsilon_{is} = \beta_1 X_{is} + \beta_2 H_{is} + \gamma_1(1 - P_{is}) + \gamma_2(1 - P_{is})X_{is} + \gamma_3(1 - P_{is})H_{is} + \varepsilon_{is} \tag{7}$$

We include the probability of admission as $(1 - P_{is})$ rather than the raw P_{is}, so the parameters β_1 and β_2 measure students' ex post preferences over program characteristics if admission were guaranteed, analogous to Equation 6.

The parameter γ_1 measures how strongly students value applying to programs with very high admission probabilities, conditional on program characteristics and peer histories. The parameters γ_2 and γ_3 measure how severely students discount program characteristics and peer histories of programs due to low admission probabilities. For example, if the elementwise signs of γ_2 are the same sign as β_1, 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 group 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 types based on rich observables using k-means clustering. In the second step, we estimate Equation 5 allowing for the error terms ε_{is} and η_{is} to have a persistent group-specific component.

Specifically, we define

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

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

In practice, this amounts to including type-by-major and type-by-university fixed effects in both stages of Equation 5. 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 type-by-program fixed effects or 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.

4.4 Identification

Identifying peer effects in the education process is a difficult task and it is not always possible to distinguish causal peer influence from other confounds 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 use high school fixed effects to control for school-level unobservables and rely only on within-high school variation in the set of peers a student interacts with, and therefore the set of peers from whom a student can learn information. We use only information transmission across cohorts of students to mitigate endogeneity concerns, such as those that might arise due to common shocks within a high school class.

Second, we use student exposure to information from peers who live in the same neighborhood but attend a different high school. In this case, we restrict our sample to neighborhoods that satisfy two conditions: first, at most 15 students per cohort, (to avoid including large neighborhoods where students are less likely to know each other), and second, students from the neighborhood attend at least two high schools. For identification, we use the set of peers from upper cohorts in the school of the neighbor.

4.5 Measuring Admission Probabilities

Estimating the role of peer influence requires a measure of the admission probabilities P_{is}. Our rich application and admissions data (Section 3.1) and Ontario's unusually transparent admissions system (Sectoin 2.2) allow us to construct an informative measure of admission probabilities, which we then treat as observable. There are two assumptions behind our estimation procedure. The first is that we can estimate valid measures of admission probabilities, i.e. that our estimates have high predictive power of true admissions. The second is that students respond to this information when building their application portfolios.

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 program unit of analysis is a major category from Table C.1 rather than a specific program, we sometimes pool programs into a single cell if a university has, for example, multiple biology 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 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.[10]

The high school GPA used for each student 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

[10]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 on course classification are available from the authors upon request.

in the outcome variable in each between-knot subset of the GPA distribution.[11]

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.

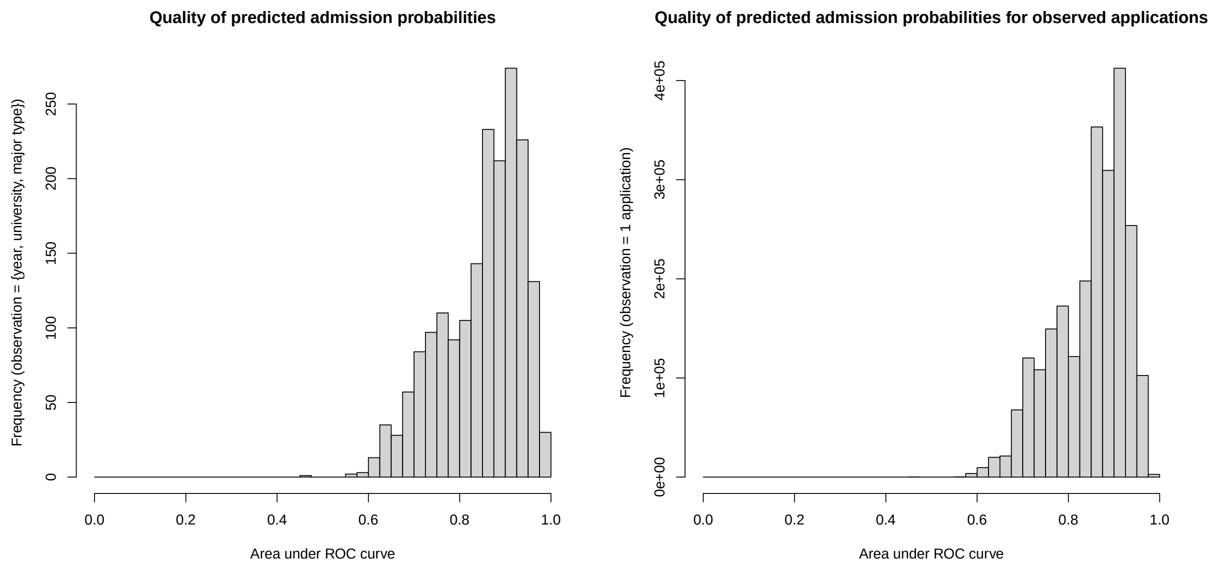
The estimated admission probability models perform well, which is unsurprising given the small gap between the information observable to the econometrician and what universities use for admissions decisions. The model fits are summarized in Figure 2. The histogram shows the distribution of area under 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% of the application sample, the AUC is fair or better (AUC ≥ 0.7), and for 72% of the application sample, the AUC is good or better (AUC ≥ 0.8).

Students' application behavior appears to respond rationally to their admissions chances. Using the admission probability estimates from Section 4.5, we can compare application portfolios by students of various calibers. If students compile their application portfolio in order to maximize total expected utility from a portfolio as a function of program qualities and probabilities of admission as in Chade and Smith (2006), then stronger students will submit more aggressive application portfolios consisting of programs that have more competitive admissions. Indeed, Figure 3 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.[12] 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

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

[12] As explained in Appendix B.1, grades above 90 are considered an A+, and grades in the 70–79 range are considered a B– to B+.

Figure 2: Performance of admission predictions



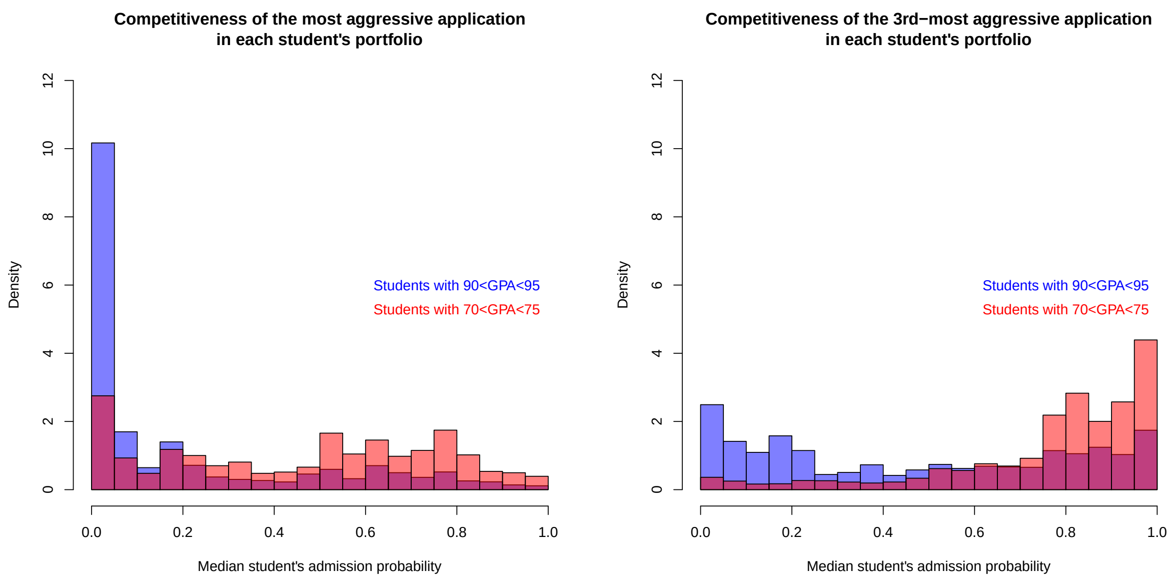
Note: Histogram of area under the ROC curve of admission probability models. The left-hand plot has the university-major-year triplet as the unit of observation. The right-hand plot has one observed application as the unit of observation, i.e. it is weighted by program popularity.

3, indicating that they apply to programs for which the median student has a lower probability of admission.

The comparisons across student qualities in Figure 3 suggest that students rationally respond to their chances of admission. We also find no evidence of outright “mistakes”. There is no evidence that students only apply to programs where they have a negligible chance of admission. The left panel of Figure 4 plots students’ own admission probabilities at the programs to which they apply, binned by ventile of student quality. It shows that the vast majority of students—all but the bottom ventile—construct application portfolios that give them a higher than 50% 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%. 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 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

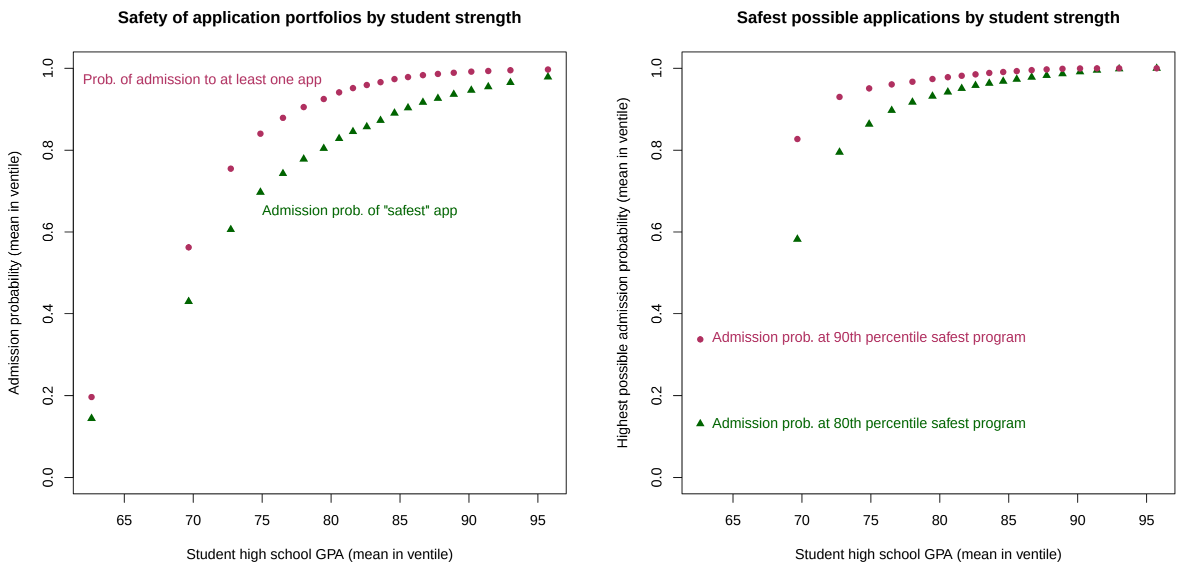
90th percentile of possible admission probabilities for students in this ventile is below 40%. Taken together, Figures 3 and 4 suggest that students rationally incorporate their chances of admission in constructing their application portfolios.

Figure 3: 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).

Figure 4: 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% 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).

5. Results

5.1 Student Types from k-means Clustering

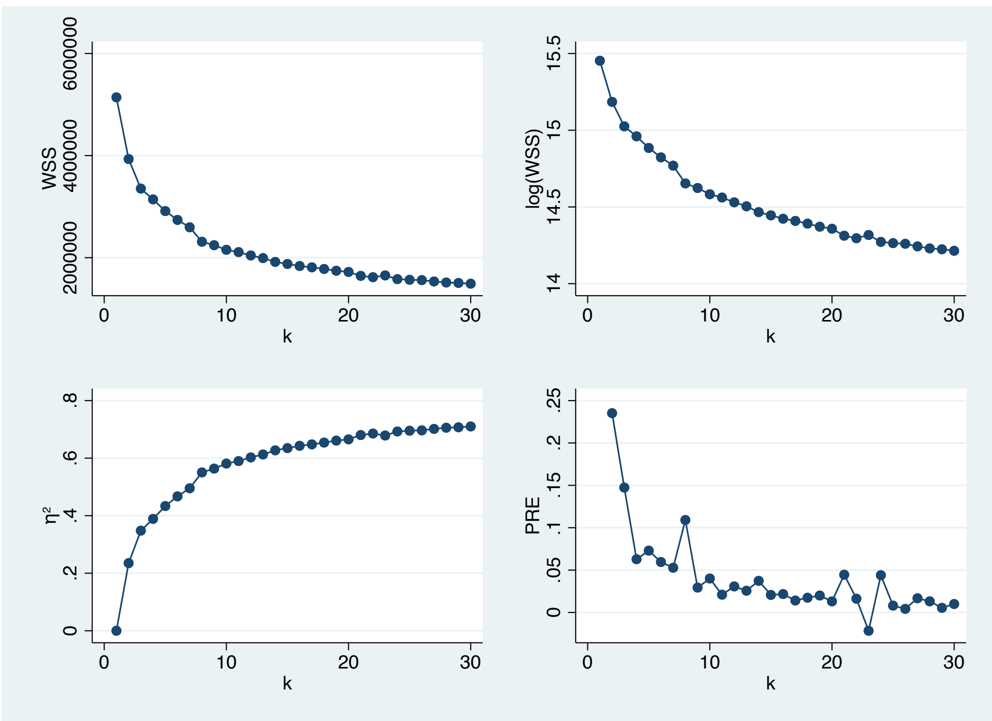
We use k-means clustering to categorize students into different types based on detailed information on courses and grades from the students course data. The objective is to capture unobserved preferences that create correlation between the errors terms in the application stage and the matriculation stage.

We use standardized measures of the following variables: average grade from top six courses, average grade from top six math classes, minimum math grade, course count of university level courses, math courses, arts courses, and social science courses. We also include indicators for taking science biology and science physics courses.

To decide the appropriate number of clusters, we use three different measures. 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).

We run the clustering algorithm for different number of clusters, from 2 to 30. Figure 5 shows the different measures for each case. The WSS graph does not show a clear kink, but overall results point to 8 or 9 clusters to be the optimal solution. The η^2 criterion points to a reduction of the WSS by 60% and PRE to a reduction of about 12% compared to $K=9$. We use $K=8$ clusters in our baseline estimates.

Figure 5: 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.

5.2 Choice Model Estimates

Table 2 shows our main estimates using both student decision stages and the k-means student clustering. The estimates reveal significant peer effects on college application decisions. The magnitude of these effects varies depending on whether the peers observed applying to those programs in previous cohorts do indeed get an offer and matriculate. 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 4.3.

Table 2: Maximum Likelihood Estimates - Neighborhood Interactions

	Raw variable		×Pr(Admit)	
(1-Adm.Prob)	0.0067	0.0169		
log(Distance)	−0.4101∗∗∗	0.0033	−0.2339∗∗∗	0.0036
Neighbor with Pr(interacted with peers who . . .)				
Applied in Cohort t-1	0.3644∗∗∗	0.0248	0.1831∗∗∗	0.0386
Applied in Cohort t-2	0.1159∗	0.0577	0.1981∗∗	0.0981
Applied in Cohort t-3	−0.1611	0.1817	0.5182	0.2636
Admitted in Cohort t-1	0.0305	0.0267	−0.1556	0.0446
Admitted in Cohort t-2	0.1493∗	0.0737	−0.6812	0.1371
Admitted in Cohort t-3	0.7131	0.3942	−5.1419	0.5076
Matriculated in Cohort t-1	0.2333∗∗∗	0.0325	−0.1438	0.0565
Matriculated in Cohort t-2	−0.1863	0.1143	0.0985	0.2026
Matriculated in Cohort t-3	0.0843	0.3106	4.1326∗∗∗	0.5105
Number of students			10,797	

Note: This table presents the parameter estimates from the maximum likelihood estimation using the neighborhood interaction 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 table shows a negative and significant coefficient on distance and its interaction with $(1 - P_{is})$, which indicates that students ex ante tend to discount programs for which their admission probability is low, and more so for programs that are farther from home. The coefficient on neighbor interactions tend to be largest for interactions with peers just one cohort above, and decay for each additional year separating cohorts. This is consistent with students interacting more with peers who are closer in age, conditional on being in the same classroom. The marginal effect of a peer who applied being admitted tends to be zero. This is consistent with admission probabilities already being observable to students. On the other hand, peers who matriculate have a large effect only for peers in the previous cohort which is also intuitive as these are the students more likely to be friends with each other. The interactions with $(1 - P_{is})$ indicate that for more selective programs, peers applications tend to have a bigger impact.

5.3 Policy Counterfactuals

Using the estimates from Table 2, we will conduct counterfactual policy simulations to evaluate the effects of changing the information environment at the college application stage.

For each student, we need to simulate the optimal portfolio under the new information environment. To do this, we follow the greedy algorithm detailed in Section 4.1. We fix the size of each student's application portfolio to the size observed in the data, as we want to isolate the effect of an information change on the identities of programs in the application set, abstracting from the decision of the number of programs to include. Peers could also influence the number of applications students decide to submit, but we focus here on the specific programs they apply holding constant the size of the portfolio.

We plan the following sets of policy counterfactuals:

1. Endowing every student with the maximum plausible amount of peer-transmitted information by setting each school to have a recent history of applications and admissions to every program.

2. Endowing every student with the minimum plausible amount of peer-transmitted information by setting histories to zero.

We will evaluate how these information treatments affect the distribution of applications across majors and across universities. For example, we will evaluate the probability that low-socioeconomic status students apply to more selective majors. We will also evaluate which types of programs and institutions are more affected. Do small, remote universities especially benefit from peer-transmitted information? Which majors become more popular with improved information, and are these majors that lead to lower- or higher-paying careers? These questions will point toward the welfare impacts of improving information.

These simulations provide insights into the potential impact of interventions aimed at improving information flow among students.

6. Conclusion

This paper shows that high school students' college application decisions vary as a function of their high school's prior history of applications, admissions, and matriculation. Our findings of information diffusion about college options across students, whether it occurs directly or through a mediator, imply that interventions targeted at one group of students can have large spillover effects for other groups. This has implications for a variety of policy issues, notably the effects of racial and socioeconomic segregation in schools.

References

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.

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.

Alvarado, S. E. and R. N. L. Turley (2012): "College-bound friends and college application choices: Heterogeneous effects for Latino and White students," *Social Science Research*, 41, 1451–1468.

Avery, C., J. S. Howell, and L. Page (2014): "A Review of the Role of College Applications on Students' Postsecondary Outcomes. Research Brief." *College Board*.

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.

Borovickova, K., A. Crema, A. Holm, and A. Nath (2023): "Peer Effects in College Applications," *Working paper*.

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

Cattan, S., K. G. Salvanes, and E. Tominey (2023): "First generation elite: the role of school networks," *Working paper*.

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

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

Ellison, G. and A. Swanson (2016): "Do schools matter for high math achievement? evidence from the american mathematics competitions," *American Economic Review*, 106, 1244–77.

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

Gaviria, A. and S. Raphael (2001): "School-based peer effects and juvenile behavior," *The Review of Economics and Statistics*, 83, 257–268.

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.

Imberman, S. A., A. D. Kugler, and B. I. Sacerdote (2012): "Katrina's children: Evidence on the structure of peer effects from hurricane evacuees," *The American Economic Review*, 102, 2048–2082.

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

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.

Sacerdote, B. (2001): "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly Journal of Economics*, 116, 681–704.

Smith, J. (2014): "The effect of college applications on enrollment," *The BE Journal of Economic Analysis & Policy*, 14, 151–188.

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.

Walton, G. M. and G. L. Cohen (2007): "A question of belonging: race, social fit, and achievement." *Journal of personality and social psychology*, 92, 82.

A. Additional Figures and Tables

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.1: Full-time Undergraduate Tuition and Fees

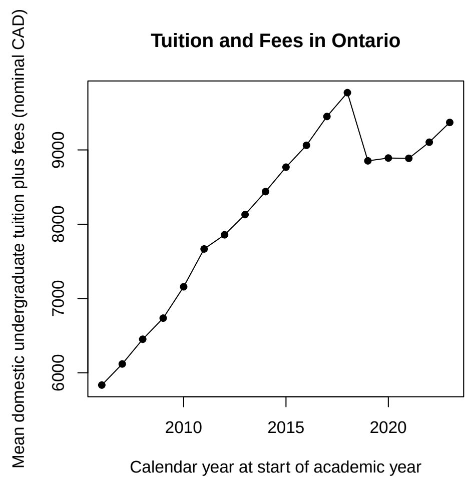
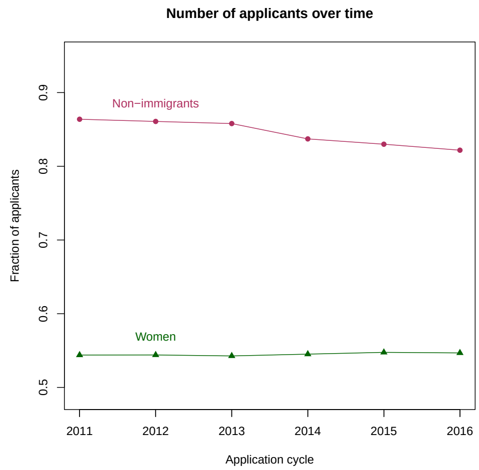


Figure plots mean tuition plus fees for Ontario undergraduates (domestic students only). Tuition is tightly regulated by the Ontario government and varies little across universities. The currency exchange rate is typically in the range of 1 CAD to 0.75–0.8 USD. Data from StatCan (2016).

Figure A.2: Applicant demographics over time



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

B. Higher Education in Ontario: Additional Context

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.[13] Figure 1 shows an example set of admission requirements, taken as a screenshot from the University of Ottawa's section of the 2011–2012 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.

[13] Unlike in the US, guidance counselors play at most a small role in steering students' college applications, as described in Appendix B.3.

Figure B.1: High School GPA Among University Applicants

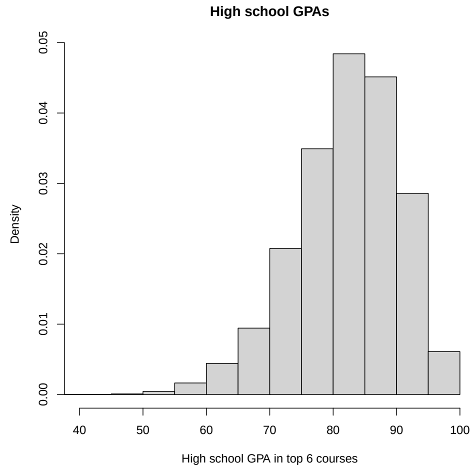


Figure plots the distribution of high school GPAs among students in our sample (those who submit at least one application to a four-year undergraduate degree in Ontario). GPA is calculated using the student's top six grade 12 courses, including English.

B.2 Grading and GPAs

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 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% of applicants). Fewer than half a dozen programs have minimum admission GPAs above 90. Even the most prestigious universities, such as Queen's University and the flagship campus of the University of Toronto, have minimum admission GPAs in the 80s for most majors.

B.3 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.4 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% 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.[14]

Table C.1 lists the categories into which we have classified majors, and the fraction of total

[14]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.

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

D. Additional Detail on Optimal Application Portfolio Algorithm

Section 4.1 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

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

a lower ex post utility of attending.[15] 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 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

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

[15] 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.

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

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