Inferring Trade-Offs in University Admissions: Evidence from Cambridge*

Debopam Bhattacharya
University of Cambridge

Julia Shvets
University of Cambridge

November, 2023

Abstract

How do elite universities balance diversity and academics during admissions, and do they face a trade-off in doing so? We develop a theory-based empirical framework to identify and quantify this potential trade-off ex-post, using post-entry outcomes. We apply our framework to admission and exam-performance data from Cambridge University. Comparing directly admitted students and second-round admits from different demographic groups yields bounds on the magnitude of trade-off, which (A) hold irrespective of whether we observe all applicant characteristics, and (B) require no information on rejected applicants. We find robust evidence of a trade-off between gender balance and performance in math-intensive subjects. It implies a weight of at least 20% on gender-diversity and at most 80% on academics in the university objective. Such trade-offs are not identified for state vis-a-vis privately-funded school students, nor for gender in equally competitive non-mathematical disciplines.

Keywords: University Admission, Affirmative action, Diversity-performance trade-off, Ex-post Evaluation, Marginal Admits, Waitlist Admission

JEL codes: D61, J71, I23, I24

1 Introduction

From political appointments to workplace recruitment and promotions, the question of diversity is receiving increasing attention around the world. An important case is demographic and so-

*Debopam Bhattacharya acknowledges financial support from the European Research Council via grant number 681565, EDWEL. The authors express deep gratitude to the Admissions Steering Committee, Dr Alexa Horner and the Central Admissions office at Cambridge for access to the data and to Renata Rabovic for her initial collaboration. We are grateful to Joe Altonji and Frank Kelly for their comments, to Amitabh Chandra and Douglas Staiger for early discussions on this topic, and seminar participants at Cornell, NYU, Stanford, and COSME gender workshop, five anonymous referees and the editor for feedback. The opinions expressed in this article are solely those of the authors and do not reflect the views of the Department of Economics, the University of Cambridge and its colleges or the European Research Council.

cioeconomic diversity in admission to elite colleges, given its potential to reduce intra- and inter generational economic inequality. For example, in the UK, a widely publicized Sutton Trust report in 2018 revealed that Oxford and Cambridge have, in recent years, recruited more students from eight specific schools than almost 3,000 other UK state schools put together. The majority of these eight schools are expensive private institutions. In the United States, many high-profile lawsuits have been contested on the issue of fair admissions cf. Regents of the University of California versus Bakke 1978 and Students for Fair Admissions, Inc. versus President and Fellows of Harvard College 2014, culminating in the US Supreme Court's 2023 decision to outlaw race-based affirmative action. At the same time, maintaining academic excellence remains a priority for such institutions. For example, the official admission statement at Oxford claims that it "seeks to admit students of the highest academic merit and potential." Similarly, Cambridge claims to "... offer admission to students of the highest intellectual potential, irrespective of social, racial, religious and financial considerations."[1]

Despite much heated exchanges in the news media and widespread public interest in the issue, rigorous statistical evaluations of existing admission practices are rare in the academic literature. In the present paper, we build an analytical framework to assess admissions ex-post by utilizing data on a post-entry outcome. In contrast to the extensive literature in economics on detecting discrimination via outcome-based tests, we ask: what objective function of the decisionmaker would rationalize the data that we observe? Our framework for answering this question consists of two elements. The first is a stylized theoretical model of diversity-performance balance in admissions, which quantifies the trade-off between the two in terms of the inter-group difference in outcomes of marginal entrants, i.e. the implicit 'cutoffs' for entry.

Because admissions are typically based on many indicators, some of which, like confidential reference letters, are typically unavailable to researchers, the marginal entrants are impossible to identify directly. So, the second element of our analysis is to then show that one can use the outcome of students from historically over-represented groups (males, privately educated etc.) entering the university from waitlists, i.e. second-round clearing, with those from under-represented groups (females, state-educated) entering directly, to bound our key parameter of interest, viz. the diversity-performance trade-off. Using waitlists for admissions is common in many, if not most, institutions. As such, our method suggests an approach that is applicable in such settings, including Cambridge from where our data come, and where the second round is known as the 'pool'.[2]

In other words, instead of presenting solely a set of reduced-form, descriptive statistics, we

[1] https://www.undergraduate.study.cam.ac.uk/applying/decisions/admissions-policy
https://www.materials.ox.ac.uk/admissions/undergraduate/admissions-criteria.html

[2] If data are available on the ranking of individual applicants, e.g. when admission is based on a single admission-test score, then it becomes possible to directly identify marginal candidates, as in Bertrand et al. 2010 and Albaek 2017. But such cases are generally less common.

provide a measure of a latent structural object that has the economic interpretation of a revealed preference parameter. This parameter can be interpreted either as an implicit weight on demographics vis-a-vis (correctly) expected future outcome, or equivalently as a measure of *systematic* deviation from correct prediction of group-specific future outcomes; based on additional data and institutional background, we make the case for the former as the more likely interpretation of our empirical results. We stress here that our theoretical model has an 'as if' character; a university is not assumed to solve exactly the optimization exercise we analyze. That is, we ask "if one analyzes the admission and ex-post outcome data as if they were generated through a trade-off exercise, then what are the implied weights on demographics vis-a-vis a future outcome of interest that would rationalize the data?"

For the empirical analysis in this paper, the specific outcome we focus on is the end-of-year exam results in the university. Although this may not capture wider implications for society of admitting students from specific backgrounds, it is a natural metric of performance to use as it is the institution's own assessment of students' academic outcomes, and hence most closely captures the university's stated objective at admissions (recall the admission statements of Oxford and Cambridge quoted above). Moreover, there is little doubt that exam results affect students' future opportunities, with employers and post-graduate programs basing their selection on them.

Note, however, that our method can be applied to *any* other post-entry outcome of interest, e.g. post-college earnings, indices of well-being during and after college, gain in academic ability from attending college, future alumni donations, etc. provided such data are available for all admitted students. This method could also potentially be applied to value-added measures of university education, as in Bleemer 2021, 2022 and Black et al. 2023, if one could simulate the counterfactual outcome for the admitted students, had they enrolled elsewhere. In the Cambridge context, to the best of our knowledge, exam results is the only outcome on which the University systematically collects data and which it uses in its own evaluations of its admission policy, and hence the empirical section of the paper focuses on those, with the understanding that our empirical conclusions could be different if a different outcome variable were used.

When applying the above methodology to the Cambridge data, our focus is on two key applicant characteristics: gender and socioeconomic background. We reach the following empirical conclusions: for gender, a diversity-performance trade-off exists in the mathematically intensive subjects, viz. Physical Sciences, Engineering, Economics and Mathematics, where the waitlisted males outperforms directly admitted females by at least 0.25 of a standard deviation in university exams. These estimates, in turn, imply that the relative weight on gender diversity in these subjects is at least 20 percent and that on academic performance at most 80 percent. Interestingly however, we find no such trade-off either for gender in other highly competitive but non-mathematical subjects, viz. Law, Biology and Medicine, nor for socioeconomic status as captured by type of school attended by the student. These findings show that the existence of a trade-off is context-dependent;

furthermore, they stand somewhat in contrast to both the long-standing public perception of Oxford and Cambridge as being socially elitist and the universities' own statements that their admissions are based solely on 'academic profile and potential' (cf. Cummings 2015).

One important and challenging question which emerges from our empirical results is *why* there is a trade-off with performance when it comes to gender diversity in mathematical subjects. We present several pieces of evidence that provide further insight into this.

Short Literature Review: Methodologically, this paper relates to the literature studying disparities in treatment of different groups, specifically to the branch that uses outcome tests inspired by Becker 1957, 1971.

The traditional approach in economics to study disparity of treatment has focussed on detecting discrimination (cf. Becker 1957, Arrow 1973, Heckman 1998). Canay et al. 2022 provide an overview of this literature, and give a formal exposition of the outcome test approach through the Roy model, highlighting important assumptions and challenges. Whilst there exists a large body of empirical research pertaining to discrimination on the basis of race and gender in the labour market (cf. Altonji and Blank 1999), law enforcement (cf. Knowles et al. 2001), credit supply (cf. Ladd 1998) and legal rulings (cf. Hull 2021); the empirical evaluation of elite university admissions is less common in comparison.[3]

There is a moderately sized literature in both economics and sociology on the consequences of positive discrimination, e.g. race-based affirmative action in US colleges. Fryer and Loury 2015 provide an overview of the broader questions surrounding the nature and impact of affirmative action. Arcidiacono et al. 2015 and Arcidiacono and Lovenheim 2016 summarize existing empirical work in economics on affirmative action in (mostly US) college admissions with a focus on race. The empirical evidence on the presence and consequences of affirmative action in college admissions is mixed. In the educational sociology literature, Boliver and coauthors, in a series of papers (cf. Boliver, 2013), have analyzed ethnicity-based inequality in UK college admissions using descriptive methods and found significant differences in admission rates by ethnic background, after controlling for observable characteristics of applicants. Bhattacharya et al. 2017 tested meritocracy of admissions at a different UK university using an assumption about the relationship between observables and unobservables, and concluded ex ante, i.e. based on pre-admission characteristics but not post-enrolment outcomes, that male applicants were being held to higher academic standards at admission. In contrast, the present paper models the preference for diversity explicitly in the objective function, does not make assumptions about unobservables as in the earlier paper, and instead uses post-enrolment academic performance to conduct an outcomes test. As such, the method outlined in the present paper can be used if a different, nonacademic outcome is the object of interest, e.g. future earnings or alumni donation. In this case, the key assumptions of the earlier

[3] Among others, papers using outcome tests to study racial discrimination in law enforcement also include Ayres 2002, Anwar and Fang 2006, Alesina and La Ferrara 2014.

paper are unlikely to hold.

As for ex-post analysis, Sander 2004 investigated race-based affirmative action in US law school admissions and concluded that it hurt the production of black lawyers, and Ayres and Brooks 2004 provided evidence to the contrary. Keith et al. 1985 concluded that affirmative action in medical school admissions led to ‘substantial integration of the medical marketplace’. On the other hand, Arcidiacono et al. 2016 found that race-based undergraduate admissions at the University of California led to lower graduation rates in STEM fields than if the beneficiaries had attended lower ranked schools, while Bleemer 2022 reported that the end of affirmative action led to worse graduation rates and lower mid-career earnings of under-represented minorities. Results similar to Bleemer were found for the US state of Texas by Black et al. 2023, whilst Bleemer 2021 also found that studying at California’s selective universities increases educational attainment and early career earnings of high achieving students from underrepresented backgrounds by the same amount or more than for their better-prepared peers.

Fu 2014 discussed a structural model of admissions which included modelling competition between colleges. In Fu’s formulation, the universities’ preference for higher tuition revenue is analogous to the preference for diversity in our case. However, Fu’s approach to modelling and identification is fully parametric. In our paper, the (set-) identification is fully nonparametric in the sense that no distributional or functional form assumption are made on unobservables and utilities.

Very recently, Arcidiacono et al 2023 have documented significantly higher admission chances for minorities with comparable test-scores as non-minorities at two highly selective US universities, while Chetty et al 2023 report that applicants from high income families are significantly more likely to attend elite institutions.

Our study is also substantively related to Autor and Scarborough 2008, who demonstrate that firms can improve diversity without sacrificing productivity by changing their selection procedures, and Kleinberg et al 2018a and Kleinberg et al. 2018b who build the case for using algorithms to improve selection in public organizations, with applications to judicial bail decisions and university admissions. In addition, Kleinberg et al 2017 show that, in such decisions, it is impossible to avoid all forms of disparate impact, so a policy that corrects for one form of discrimination would introduce another. Our analysis illustrates this: in section 3.1.1, we show graphically that, theoretically, it is generally impossible to maintain both equal success rates across genders (or equal gender shares among admitted) and equal thresholds at admissions. Then subsequently, we show empirically that the university has managed to equalize success rates across genders in the STEM fields, but not the admission thresholds.

The rest of the paper is organized as follows: In Section 2, we describe the context and admission process at the University of Cambridge. In Section 3, we first offer a simple theoretical model of decisions by a university which values both performance and diversity. We then develop our econometric approach, using the two-tier nature of admissions. In Section 4, we describe our data. In Section 5, we show the results obtained by applying our methods to the Cambridge data, focussing on gender. In Section 6, we present the results on socioeconomic background. Section 7 summarizes and concludes. The Appendix presents additional results and further robustness checks.

2 Empirical context

This section describes our empirical setting, that of the University of Cambridge. We first provide some discussion on the dual objectives of performance and diversity faced by the University, then how performance is assessed, and finally the details of how students are admitted. These inform our theoretical model and empirical identification, as presented in Section 3.

All undergraduate students and most academics in Cambridge belong to one of its 29 colleges[4]. Colleges, in addition to providing accommodation and meals, fulfil two crucial functions. First, they deliver nearly all small group teaching to undergraduates, so called supervisions. Second, they make undergraduate admission decisions on behalf of the University (more on this is in section 2.3).

2.1 University objectives and diversity concerns

The University of Cambridge is one of the oldest and most prestigious higher education institutions in the world. Every year, around 20,000 students apply for 3,500 places in Cambridge's undergraduate programme, making it among the most selective, internationally.

The objectives of the University's admissions are articulated in its Admissions Policy (University of Cambridge, 2023):

> The principal aim of the Admissions Policy of the Colleges of the University of Cambridge is to offer admission to students of the highest intellectual potential, irrespective of social, racial, religious and financial considerations.
>
> Two further aims are:

- aspiration – to encourage applications from groups that are, at present, under-represented in Cambridge,
- fairness – to ensure that each applicant is individually assessed, without partiality or bias, in accordance with the policy on Equal Opportunities, and to ensure that, as far as possible, an applicant's chance of admission to Cambridge does not depend on choice of College.

Based on this, we make three observations. First, when we model the University's admissions, 'offering admission to students of the highest intellectual potential' should form the main part of

[4]There are 31 colleges in Cambridge, but two only admit postgraduate students.

the objective function. Second, in 'further aims' the University's raises concerns about under-represented groups, and so we want to also allow for this when modelling its objective. Indeed, the statement talks about 'encouraging applications' from under-represented groups. In our model, we allow for the possibility that the university might care directly about the number of students it admits from such groups; we then let the data tell us whether this is the part of the objective function or not. This also reflects the fact that diversity in higher education is a key concern at most leading universities around the world, and not just at Cambridge. Third, other than fair treatment of applicants, there are no further aims articulated in the University's statement. This is important, since our identification is based on the assumption that the university is not pursuing goals other than those included in the model's objective function, as is also discussed in Canay et al. 2022.

2.1.1 Diversity concerns

We focus on two student characteristics which are the key focus of diversity concerns, both worldwide and specifically for the UK, viz. gender and socioeconomic background.

Gender: Although the overall enrolment of girls often equals or exceeds that of boys, women's participation in STEM disciplines has been an area of concern around the world. As we show later, this is also a concern at Cambridge, where women are significantly underrepresented in these subjects (see Section 4).

Socioeconomic background: Most universities around the world are concerned about expanding participation from less advantaged socioeconomic groups. This is of particular concern in the UK, where universities including Cambridge, are primarily funded by tax revenues. One of the key measures used by the UK government to assess the socioeconomic background of students is the type of school attended prior to university. For this purpose, all UK schools are divided into two groups: state-funded schools that are free to attend, and independent, fee-charging schools, which we will refer to as privately-funded.[5] The latter typically enrol children from households of higher socioeconomic status, and these children tend to be over-represented at the top UK Universities relative to their proportion in the population. One of the key widening participation targets that the UK government sets for the universities pertains to the proportion of students from UK state-funded schools. Currently, the government's target for Cambridge University is 64% and is set to increase to 69% by 2024 (University of Cambridge 2018, 2020). The type of school attended prior to university is therefore a natural metric for a university keen on enhancing socioeconomic diversity on campus.

Going forward, we take females to be the 'protected' group for gender and the state-educated

[5] In total, about 6% of UK children attend privately-funded schools (Sibieta 2021), although this proportion rises for 16+ year olds.

for socioeconomic status. Our subsequent analysis will allow us to verify whether indeed these are protected groups, i.e. whether the university is willing to sacrifice expected academic performance in order to admit more students from them.

2.2 Cambridge exams as outcomes

Recall that Cambridge states that selecting 'students of the highest intellectual potential' is the main objective of its admission policy. To take our model to the data, we need a measure of student outcomes that closely correspond to this objective. Using the students' subsequent exam results seems to be a natural choice, since it is the outcome by which the University itself measures its objectives and it carries important consequences for the students in terms of future opportunities, as discussed in the introduction.

Most undergraduate degrees in Cambridge are three years long.[6] Every year, students enrol in multiple courses. For each course, all students sit the same centrally-set exams, with no variation by college, and their scripts are marked blindly, making their scores comparable with one another.[7] These exams are typically sat at the end of the year. The final transcript at the end of the degree lists the student's performance in each year separately.[8]

2.3 Admission procedure

In contrast to the liberal arts approach in countries like the United States, the UK university system including Cambridge, requires prospective students to apply for a specific subject, e.g. Mathematics, Law, Engineering etc. Having chosen the subject they want to study for their degree, a student applies to one of Cambridge's 29 constituent colleges, which take undergraduate students. The student can only apply to one subject and one college. The college conducts assessment for that applicant in that subject and makes the final decision on whether to admit the applicant. Each college has an approximately fixed number of places for each subject, which is agreed at the university level. After enrolling, changing subject is difficult, though not impossible. Once admitted to a specific college, students cannot change to a different one. For the purpose of this paper, we view the aggregate admission process as pertaining to the entire university. This is similar to other institutions, where different admission officers make decisions about individual applications but the general policy is set at the university level and so admission results are typically assessed at the level of the university as a whole. Indeed, crime-detection, legal sentencing etc. also have

[6] In our data, notable exceptions are Engineering, Mathematics and Medicine that (can) last longer. We discuss this in Section 4 where we describe our data.

[7] When marking exams, faculty members do not know the names of the candidates, and only see their registration number on the script.

[8] During the period of our dataset, there was no aggregation across years to produce a final indicator, unlike an aggregate GPA in US universities.

this general feature where individual police officers and individual judges respectively make the relevant decisions, and it is the aggregate decision of the institution as a whole that is assessed by researchers.

The exact admission procedure at Cambridge varies slightly by subject and college, but the general procedure consists of the following steps. Students in most subjects take an admission test, which is the same across all colleges. Those who perform well are then invited to a second assessment – usually an interview – and the rest are rejected. After the second assessment, the college makes its admission decisions, based on the results of assessments and the strength of the student's file, which includes school-leaving exam scores, reference letters and the statement of purpose.

The admission decision has two tiers. The students deemed strongest are admitted directly by the college they had applied to. We will refer to these first-tier admissions as 'direct admission'. The college also identifies a subset of the remaining students who did not make it into the first group but are still relatively strong, and they are placed in a common 'pool'. Those colleges that have not filled all of their places with the first-tier candidates admit one or more students from this pool.[9] We refer to the candidates admitted in this way as the pooled or second-tier admits. About 20% of students in our sample are admitted via the second tier. This two-tier structure of the admission system plays an important role in our empirical strategy.

Cambridge is seen as a top UK University and one of the best in the world for undergraduate education. Hence, UK candidates who get an admissions offer at Cambridge almost never turn Cambridge down, and the fraction of foreign candidates who do so is small. So, by and large, admission offers are made to the best students, without much regard to whether they may be considering a different university.[10]Cambridge offers are made early (in January for October) and so are typically conditional on attaining certain grades in high school or other specific exams taken later. A minority of candidates who get an offer do not meet their conditions, and they are eventually rejected by the University (see Appendix C.1 for details and Figure 5 for the summary of Cambridge's admission process.).

3 Formal framework and identification

This section sets up an empirically motivated theoretical framework to identify and quantify the potential diversity-performance-trade-off accepted by the university, which we then take to the data in Section 5.

Toward that end, we first propose a model of the university's admission (section 3.1) which de-

[9]The pool takes place soon after the first tier decisions are made, and, after its completion, the offers for first and second tier decisions are sent out on the same day.

[10]Rules governing UK applications do not allow students to apply to Cambridge and Oxford at the same time.

rives the marginal condition indicating the trade-off. Underlying our methodology is Becker's 1971 insight that if preferences for diversity cause decisionmakers to deviate from the aim of maximizing an eventual outcome, then, at the margin of a positive decision, those who are favoured by the policy would have a lower expected outcome.

Second, taking the conditions derived in the model directly to the data is challenging due to unobservability of a number of factors known to the admission decisionmakers, making it impossible to identify marginal candidates. To address this infra-marginality problem, we develop an identification strategy based on the two-tier admission process followed in Cambridge (section 3.2).

3.1 Model of trade-off

We first provide some graphical intuition to motivate our formal model. Then we explain our main assumptions, before developing the model.

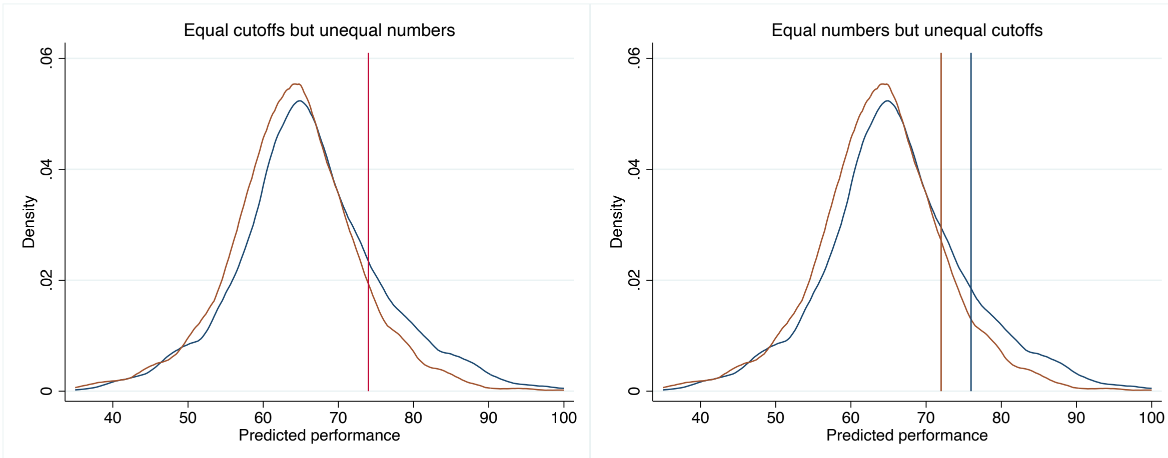
3.1.1 Graphical intuition

Consider Figure 1 which plots the distribution of expected value of the outcome of interest (for example, post-entry academic performance predicted from pre-admission credentials) for two groups of applicants, termed Red and Blue. The left graph corresponds to the admission decision which uses the same cutoff for expected outcome in both groups, which is equivalent to equalizing the expected outcome of marginal candidates from the two groups.[11] This, however, leads to different numbers of students entering from each group, because the right tail of the Blue distribution is thicker than that of the Red distribution. This corresponds to the case where the decisionmaker wants to simply admit students with the highest expected outcome and thus ends up admitting more Blues than Reds. The right graph, on the other hand, corresponds to the decision where the number of Blue and Red students entering is made equal by fixing the admission cutoff to be lower for the Red group. These represent the two extreme cases where the decisionmaker cares only about expected outcome or only about equality in numbers.

Now consider a decisionmaker who values both objectives. We say that a trade-off between demographics and expected outcome exists if such a decisionmaker has to sacrifice some amount of expected outcome in order to improve demographics and vice versa. This is the case that we formalize and estimate using the data.

[11] In what follows, we will use the terms 'difference in cutoffs' and 'difference in marginal performance' interchangeably.

Figure 1: Potential diversity-performance trade-off



3.1.2 Key assumptions

The decisionmaker's problem rests on three key assumption, viz. the objective, the information and the constraints.

The decisionmaker's objective has two parts. First, we assume that the university aims to maximize the expected value of a specific outcome across students, academic performance in our context (see introduction and Section 2.2). In addition, the university is assumed to potentially care about diversity in its objective function, by putting a weight on the number of students it admits from a currently underrepresented group. This reflects widespread concerns about diversity in higher education in general and in Cambridge specifically. Section 2.1 discusses how these assumptions emerge from our institutional context. In the main model we assume that the decisionmaker is risk-neutral, and then relax this assumption in Section 3.3.

Second, we assume that during admissions the university observes the applicants' characteristics, which it uses to make an unbiased assessment of their future performance. In Section 3.4, we relax this and discuss implications of biased beliefs.

Third, we assume that the university faces a constraint on the total number of students it can admit, which reflects UK government regulations and the university's own resource constraints.

3.1.3 Model and analysis

We model the first round admission decision of the university, i.e. which applicants to admit directly from two groups denoted by $G = g$ and $G = h$. Let h be the 'protected' category – e.g. g can be males and h females, or h can be state-school educated and g privately educated, etc. Suppose that the number of applicants in the two groups g and h is N_g, N_h respectively and total number of places is M. Let Y denote the random variable representing the potential outcome of interest (e.g. future academic performance). Assume that as its efficiency objective, the university wants to maximize aggregate future outcome, and as its diversity objective, it wants to increase the number of students from the 'protected' group h.[12]

The university observes characteristics X but not Y when making the admission decision. The researcher observes a subset of X for all applicants and Y for admitted applicants. We assume for now that the university makes an unbiased prediction for performance given observed characteristics.

Accordingly, suppose the university's overall objective is to decide which subset of applicants defined by values of X should be admitted, i.e. pick the sets $\mathcal{X}_g \sqsubseteq support(X|G=g)$ and $\mathcal{X}_h \sqsubseteq support(X|G=h)$ to satisfy the following objective:

$$\max_{\mathcal{X}_g,\mathcal{X}_h} \left[N_g E\left(Y \times 1\left\{X \in \mathcal{X}_g\right\}|G=g\right) + N_h E\left(Y \times 1\left\{X \in \mathcal{X}_h\right\}|G=h\right) + \beta N_h \Pr\left\{X \in \mathcal{X}_h | G=h\right\}\right], \tag{1}$$

where $\beta \in \mathbb{R}$, subject to

$$N_g \Pr\left\{X \in \mathcal{X}_g | G=g\right\} + N_h \Pr\left\{X \in \mathcal{X}_h | G=h\right\} = M. \tag{2}$$

Here, the first part of the objective function $N_g E\left(Y \times 1\left\{X \in \mathcal{X}_g\right\}|G=g\right) + N_h E\left(Y \times 1\left\{X \in \mathcal{X}_h\right\}|G=h\right)$ equals the aggregate correctly expected (by the university) future outcome of the admitted g-types plus that of the admitted h-types; the second part $N_h \Pr\left\{X \in \mathcal{X}_h | G=h\right\}$ is the number of h-types admitted, while β is the relative weight on diversity vis-a-vis outcomes. The constraint (17) simply equates the number of admits to the number of available places. We will use our data to learn the value of β which will throw light on the extent of trade-off that the admission process entails. Note that if $\beta > 0$, then the university values applicants from group h, if $\beta < 0$, it has an aversion to them, and if $\beta = 0$, then it is neutral.

Note that the objective function in (16) implicitly makes the simplifying assumption that the distributions of post-entry outcomes are not affected by the fraction of g or h types admitted, i.e. rules out peer-effects of the type found in Bostwick and Weinberg 2022. Indeed, when we test for presence of peer-effects, we do not find them (see Section 5.6.1).

To make further progress, define the random variables $A_g(X) \equiv E(Y|X, G=g)$ and $A_h(X) \equiv E(Y|X, G=h)$. They denote expected values of Y as inferred by the admission officers on the basis of the X's, respectively. As the value of X varies among g/h-type applicants, so does the

[12]For simplicity, we abstract from the possibility that some of the applicants to whom the university makes an offer may not arrive. Allowing for this does not change the main implication of the model, summarized by Claim 2, even if the rates of arrival are different across groups (see discussion in Appendix C.1.)

value of A_g/A_h. Let F_g, F_h denote the marginal CDFs and f_g, f_h denote marginal densities of A_g, A_h, respectively.

We now state and prove our main theoretical results which provide a solution to the problem (16) in terms of quantities potentially (set-)identifiable from the data.

Claim 1 *The solution to the problem (16) solves the problem:*

$$\begin{array}{l} \max_{g_1,h_1}\left[N_g\int_{g_1}^{\infty}af_g\left(a\right)da+N_h\int_{h_1}^{\infty}af_h\left(a\right)da+\beta N_h\left(1-F_h\left(h_1\right)\right)\right] \\ s.t.\ N_g\left(1-F_g\left(g_1\right)\right)+N_h\left(1-F_h\left(h_1\right)\right)=M \end{array} \tag{3}$$

for some real numbers g_1 and h_1 (Proof in Appendix)

Intuitively, this means the university's problem is that of choosing the thresholds g_1 and h_1 such that, in each group, all applicants with expected performance above these cutoffs are directly admitted. In (3), the term $\int_{g_1}^{\infty}af_g\left(a\right)da$ equals the expected future outcome of admitted g-types, $\left(1-F_g\left(g_1\right)\right)$ is the fraction of g-type applicants who are admitted, and analogously for h-types.

We now show that the solution to the problem (3) takes a particularly simple form that relates the parameter β to the difference in admission cutoffs for the two groups.

Claim 2 *If A_g and A_h are continuously distributed, then the problem (3) has a unique interior maximum, where $\beta=g_1-h_1$. (Proof in Appendix)*

The expression g_1-h_1, the difference between the cutoffs for direct admission, is the difference between the expected performance of marginal candidates from the two groups. If the decisionmaker were maximizing just the performance objective ($\beta\to 0$), they would set $g_1-h_1=0$. Conversely, if they were only maximizing h-type enrolment ($\beta\to\infty$), they would set $N_h s=M$. Our goal is to learn the value of β, or equivalently g_1-h_1, from the data.

It is important to note that the proof of claim 2 does not require any functional form assumptions on unobservables, e.g. that expected performance is normally distributed etc. In that sense, it is a nonparametric (identification) result that expresses the latent structural object β characterizing preferences, in terms of the reduced form parameter g_1-h_1, using an economic model of constrained optimization.

3.1.4 Roy model and treatment assignment

Problem 16 can be interpreted in the light of an Extended Roy Model (ERM), as in Canay et al. 2022, who develop a general framework for thinking about outcome tests of discrimination. Our Claim 2 is analogous to their theorem 4.2, but differs from it owing to the binding capacity constraint. In Appendix B, we elaborate further on this analogy.

Problem 16 is also similar in spirit to that of optimal treatment assignment, subject to resource constraints, considered previously in Bhattacharya and Dupas 2012. That paper was concerned with recommending the optimal decision based on results from a randomized experiment, as opposed to assessing a decisionmaker's preferences from an observational dataset, which is our goal here.

3.2 Identification

Suppose g_1 and h_1 were observed empirically. Then, bringing the model to the data would allow us to do three things:

1. We can identify whether there is a diversity-performance trade-off. As implied by the derivations above, the decisionmaker makes a trade-off if $g_1 \neq h_1$, and does not if $g_1 = h_1$. n the latter case, when we cannot reject $g_1 = h_1$, the likely explanation is that an acceptable number of h-types get in, so that the decisionmaker does not need to consider demographics separately in admissions, and they set $\beta = 0$.

2. Where the decisionmaker makes a trade-off, we can find the weight β they place on the number of admits from the protected group h.

3. We can crudely measure the size of the trade-off ex-post, in terms of actual performance forgone and additional students admitted from group h (see section 5.3).

How does one measure g_1 and h_1 empirically? If admission depended on a single test-score, as for example in some engineering colleges in India (cf. Bernard et al. 2010), the values of g_1 and h_1 would be known from the lowest admission test-scores among the admits of each group g and h. But in most cases, including at Cambridge, admission depends on a lot of different variables associated with each applicant, some of which are typically unobservable to the researcher. Therefore, it is not possible in general to learn the values of g_1 and h_1 directly from the data. To address this, we develop a method to identify the differences in admission cutoffs, by exploiting the fact that many students enter universities, including Cambridge, via waitlists or second round clearing.

Toward that end, define $\varepsilon_g = Y - E(Y|X, G = g)$, implying by definition that $E(\varepsilon_g|X, G = g) = 0$, and therefore $E(\varepsilon_g|A_g) = 0$, since $A_g = E(Y|X, G = g)$ is solely a function of X. In particular, this implies that with $F_g(\cdot)$ denoting the marginal CDF of A_g, any set C with $F_g(\cdot)$-positive probability, we have that:

$$E(\varepsilon_g|A_g \in C) = \int_{a \in C} \underbrace{E(\varepsilon_g|A_g = a)}_{=0} dF_g(a) = 0. \tag{4}$$

This implication will be used below.

Recall that in our setting the university practices a two round admission process, first is direct admission, and second is the so-called pool. Hence, the admission decision can be summarized via three cutoffs $g_2 < g_3 < g_1$ such that if $A_g > g_1$, then the applicant is admitted directly, if $A_g < g_2$ then s/he is rejected straight away, and if $g_2 < A_g < g_1$, then the candidate is put in the pool. Finally, if $g_3 < A_g < g_1$, then the candidate is eventually admitted from pool. For h-type applicants, denote the analogous quantities by A_h and h_1, h_2, h_3 respectively.

3.2.1 Bounding β

We will now show how to identify lower and upper bounds of $\beta = g_1 - h_1$.[13] The econometrician observes the outcomes of all entrants and, in particular, those of pooled g-type admits $(Y|g_3 < A_g < g_1, G = g)$, and of directly admitted h-type admits $(Y|A_h \geq h_1, G = h)$. Therefore, the average outcome of pooled g-type admits equals

$$
\begin{aligned}
&E[Y|g_3 < A_g < g_1, G = g] \\
&= E[A_g + \varepsilon_g|g_3 < A_g < g_1, G = g] \qquad (5)\\
&= E[A_g|g_3 < A_g < g_1, G = g] + \underbrace{E[\varepsilon_g|g_3 < A_g < g_1, G = g]}_{=0 \text{ by } (4)} \\
&= E[A_g|g_3 < A_g < g_1, G = g] < g_1, \qquad (6)
\end{aligned}
$$

while the average outcome of directly admitted h-type admits equals

$$
\begin{aligned}
E[Y|A_h \geq h_1, G = h] &= E[A_h + \varepsilon_h|A_h \geq h_1, G = h] \\
&= E[A_h|A_h \geq h_1, G = h] + \underbrace{E[\varepsilon_h|A_h \geq h_1, G = h]}_{=0 \text{ by } (4)} \\
&= E[A_h|A_h \geq h_1, G = h] \geq h_1. \qquad (7)
\end{aligned}
$$

It follows from (6) and (7) that

$$g_1 - h_1 > E[Y|g_3 < A_g < g_1, G = g] - E[Y|A_h \geq h_1, G = h]. \qquad (8)$$

The RHS, which is estimable from our data, thus provides a *lower bound* on the difference in cutoffs for direct admissions $g_1 - h_1$. In particular, if the average outcome for pooled g-type admits is (weakly) higher than that of directly admitted h-type admits, i.e.

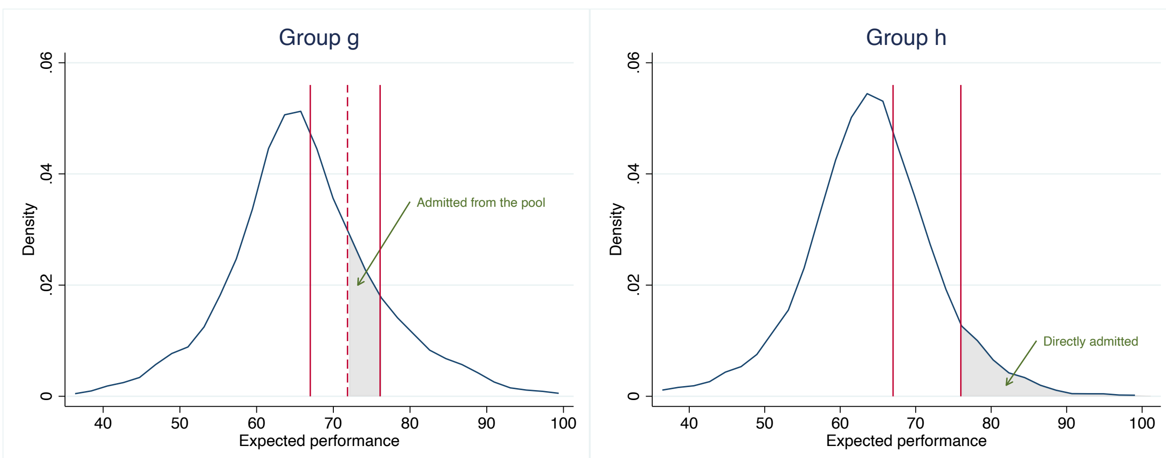
$$E[A_g + \varepsilon_g|g_3 < A_g < g_1, G = g] \geq E[A_h + \varepsilon_h|A_h \geq h_1, G = h], \qquad (9)$$

then $g_1 > h_1$. In fact, if $\Pr(g_3 < A_g < g_1|G = g) > 0$, and $\Pr(A_h > h_1|G = h) > 0$ – corresponding to the likely scenario that expected performance is continuously distributed – then even *equality* of mean outcomes, i.e. $E[A_g + \varepsilon_g|g_3 < A_g < g_1, G = g] = E[A_h + \varepsilon_h|A_h \geq h_1, G = h]$ will also imply that $g_1 > h_1$.

[13] Indeed, $g_3 - h_3$ is also an interesting parameter of interest, but its identification is impossible because of standard inframarginality problems, as noted by a referee.

These ideas are graphically illustrated in Figure 2, where the distribution of conditional expected outcome of the two hypothetical groups are plotted in the two graphs, and the rightmost vertical lines in the left and right graphs represent g_1 and h_1, respectively. The shaded areas in the left and right graphs correspond to sets $(A_g + \varepsilon_g)|g_3 < A_g < g_1, G = g$ and $(A_h + \varepsilon_h)|A_h \geq h_1, G = h$ respectively. If $g_1 = h_1$, then *every* individual in the shaded area on the left (for group g) must score lower than *every* individual in the right shaded area (for group h), up to random noise. Therefore, the average on the left must be less than the average on the right. A contradiction implies that the rightmost vertical line in the left figure must be at a higher level of ability than the rightmost vertical line in the right figure. That is the basis of our (partial) identification approach.

Figure 2: Identifying differences in cutoffs



Remark 1 *Note that reversal of the inequality (9), i.e.* $E[A_g + \varepsilon_g|g_3 < A_g < g_1, G = g] < E[A_h + \varepsilon_h|A_h \geq h_1, G =$ *is consistent with both* $g_1 > h_1$ *and* $g_1 < h_1$*, and is therefore inconclusive.*

A completely analogous argument leading to (8) applies mutatis mutandis to the *upper bound*, viz.

$$g_1 - h_1 < E[A_g + \varepsilon_g|A_g > g_1, G = g] - E[A_h + \varepsilon_h|h_3 < A_h < h_1, G = h] \tag{10}$$

The RHS of (10) is the difference in average performance between directly admitted g-types and the h-types admitted from the pool, and it provides an estimate of the upper bound of the difference in admission cutoffs.

3.2.2 Identifying the trade-off

As discussed in Section 3.1, to identify a diversity-performance trade-off we need to know whether $g_1 - h_1 > 0$. Since we do not observe $g_1 - h_1$, but rather its lower and upper bound, we have some freedom to decide when the trade-off is present. We will use the most conservative rule and say that we 'identify a trade-off' only when *both* estimated lower and upper bounds of $g_1 - h_1$ are significantly above zero. Then, RHS of (8) and RHS of (10) give us the lower and the upper bound for the magnitudes of the trade-off at the margin. We use them to bound β, the relative weight on the demographic characteristic in the decisionmaker's objective.

In contrast, when the lower bound of the cutoff difference is significantly below zero or includes zero in its confidence interval and the upper bound is significantly above zero, we will say that there is 'no conclusive evidence of a trade-off.[14]

3.3 Relaxing assumptions: Risk aversion

We now relax the assumption that decisionmaker is risk neutral, and instead suppose they are risk averse and base the admission decision on $B_g \equiv E\left(U\left(Y\right)|X, G = g\right)$ for a concave, increasing $U\left(\cdot\right)$, instead of $E\left(Y|X, G = g\right)$. Now if the distribution of $Y|g_3 < B_g < g_1, G = g$ first-order stochastic dominates (FOSD) that of $Y|B_h \geq h_1, G = h$ (as found in the empirics and reported in Section 5.2), then $E\left[U\left(Y\right)|g_3 < B_g < g_1, G = g\right] \geq E\left[U\left(Y\right)|B_h \geq h_1, G = h\right]$ for *all* increasing $U\left(\cdot\right)$, and therefore,

$$g_1 - h_1 > E\left[U\left(Y\right)|g_3 < B_g < g_1, G = g\right] - E\left[U\left(Y\right)|B_h \geq h_1, G = h\right] \geq 0. \tag{11}$$

Thus under FOSD, the conclusion of higher admission cutoffs for g-types, i.e. the positive sign of $g_1 - h_1$ is robust to risk-aversion considerations.

3.4 Relaxing assumptions: Wrong beliefs

Our model in Section 3.1 assumes that the decisionmaker forms correct beliefs about expected outcomes of admitted students. So any differences in the performance of marginal candidates are due to the university's desire to diversify student intake by setting different admission cutoffs. However, if the decisionmaker has systematically incorrect beliefs, they can lead to differences in performance of marginal candidates, even in absence of the diversity motive. In this case, the university unintentionally sets admission cutoffs at different levels due to incorrect beliefs. A number of authors, including Bohren et al. 2021 and Canay et al. 2022, have previously noted this type of observational equivalence. In this section we consider three examples of plausible wrong beliefs that might give rise to this.

[14] The remaining possibility is RHS of (8) > 0 and RHS of (10) is ≤ 0, i.e. the candidates taken from the pool outperform directly admitted ones. This would imply that the two-tier admission system is not working as intended, invalidating our identification strategy. In Section 4.3 we show that this never occurs in our data, i.e. both in aggregate and within each group, pooled candidates underperform directly admitted ones.

3.4.1 Biased beliefs favouring h

The university may have biased beliefs favouring h-type, predicting that each h-type person will score β more than what they truly would.

Then, even if the university's objective is purely to maximize the expected performance, the consequences of these incorrect beliefs will be that the marginal g-type will perform better than the marginal h-type, implying $g_1 - h_1 = \beta$, same as in Claim 2. This can be seen by rewriting our objective function in (3) as

$$\max_{g_1,h_1}\left[N_g\int_{g_1}^{\infty} af_g\left(a\right)da + N_h\int_{h_1}^{\infty}\left(a+\beta\right)f_h\left(a\right)da\right],$$

where the CDF of the true future outcomes of group h is shifted to the right by β. Hence, the empirical prediction of this is the same as when the university cares about diversity but is unbiased.

3.4.2 Ignoring group differences

Another alternative is that group identity is in fact *ignored* by admission officers during decision-making.[15] This can produce systematically biased beliefs as well, e.g. if given the same observed covariates X, the g-types have higher predicted outcomes, i.e., $E\left(Y|X=x,g\right) = E\left(Y|X=x,h\right) + \beta$. However, the admission officers do not take group differences into account, and instead use $E\left(Y|X=x\right)$ to make the decision, i.e. admit if $E\left(Y|X=x\right) \geq \gamma$, for some threshold γ. Now, since

$$\begin{aligned} E\left(Y|X=x\right) &= \frac{N_g}{N_g+N_h}E\left(Y|X=x,g\right) + \frac{N_h}{N_g+N_h}E\left(Y|X=x,h\right) \\ &= \frac{N_g}{N_g+N_h}\beta + E\left(Y|X=x,h\right) \\ &= E\left(Y|X=x,g\right) - \frac{N_h}{N_g+N_h}\beta, \end{aligned}$$

it follows that for those deemed to be the marginal entrants in the two groups satisfy

$$\begin{aligned} E\left(Y|X=x,h,E\left(Y|X=x\right)=\gamma\right) &= \gamma - \frac{N_g}{N_g+N_h}\beta, \\ E\left(Y|X=x,g,E\left(Y|X=x\right)=\gamma\right) &= \gamma + \frac{N_h}{N_g+N_h}\beta; \end{aligned}$$

and so the difference between the two equals

$$\gamma + \frac{N_h}{N_g+N_h}\beta - \gamma + \frac{N_g}{N_g+N_h}\beta = \beta.$$

This situation could arise when the 'returns' to applicant characteristics X observable during admissions differ by group, but the admission decisionmakers do not take this into account (e.g.

[15] We are grateful to an anonymous referee for raising this possibility.

as in Kleinberg et al. 2018b). Whilst the main prediction of this would be the same as when the unbiased decisionmaker cares about diversity, this mechanism generates two more empirical predictions: (i) the returns to pre-admission characteristics are different across the two groups and (ii) these differences help explain the observed differences in thresholds. In section 5.6.2, we show that, when it comes to gender, there is some evidence of (i) but not of (ii). These results suggest that admission decisionmakers ignoring group differences is unlikely to explain the differences in the thresholds which we observe.

3.4.3 Incorrect weights on pre-admission qualifications

Another possibility is that admission officers place the wrong weights on different pre-admission qualifications. This may lead to the same prediction as under the unbiased but diversity-concerned decisionmaker, i.e. to differences in observed admission thresholds, if these mistakes systematically benefit one of the groups more than the other.[16] In Section 5.6.2 we report two findings: first, admission officers appear to underweight more specialist mathematical training and overweight more general training of applicants, relative to their importance for student outcomes. Second, female applicants have less of the former and more of the latter training compared to male applicants. Although these findings are only suggestive, they are consistent with such incorrect beliefs.

To summarize, by purely looking at marginal candidates, it is not possible to say whether diversity motives or wrong beliefs drive the observed differences in admission cutoffs. Further analysis with additional data lends some support to incorrect weights on qualifications (3.4.3) but not to the neglect of group differences (3.4.2). Additionally, the idea that the university simply systematically overstates future performance of females (3.4.1) is somewhat implausible since the gender-gap in exam attainment is well documented and widely discussed at Cambridge, cf. Ingrey 2021.

At the same time, the fact that we find conclusive evidence of lower admission cutoffs for females only in mathematically intensive subjects where females are in the minority, but not in other subjects, where they constitute 50% or more (section 5), lends support to the diversity mechanism. The fact that the university maintains similar offer probabilities across genders in most subjects, with females facing an offer probability that slightly but statistically significantly higher when they are in a minority (MI subjects) and slightly but significantly lower when they are in the majority (non-MI subjects) also points at the diversity motive (see Tables 2 and Appendix D). So in what follows, we will continue to use the diversity model for interpretation, whilst keeping in mind that one cannot fully rule out systematically biased beliefs.

[16] We are grateful to an anonymous referee for suggesting this.

4 Data

In this study, we utilize administrative micro-data on individual students from the Cambridge Admissions Office. The data contain student characteristics: gender, school-type, pre-admission qualifications, subject, college, whether admitted directly or from the pool, and exam performance after entering Cambridge. table Cambridge offers undergraduate degrees in roughly 30 subjects. In this paper, we focus on seven larger subjects which jointly account for about half of Cambridge's undergraduates. These are Economics, Engineering, Mathematics, Biological Sciences, Law, Medicine and Physical Sciences.

Among these, Economics, Engineering, Mathematics and Physical Sciences have higher quantitative content and require more advanced mathematical preparation than Law, Medicine and Biological Sciences. In Cambridge, the former group of subjects have a minimum mathematics requirement that the applicants must fulfil before coming to Cambridge, while the latter group do not (cf. University of Cambridge 2022a).[17] Henceforth, we will refer to the former group as mathematically intensive (MI) subjects. The two groups of subjects are similarly competitive as measured by the offer probabilities: they are 22% for non-MI subjects and 24% for MI subjects.

Our sample consists of just under 6 thousand students, representing all students who entered Cambridge in 2013-2016 to study these subjects, and stayed till the end of their degree. In Appendix C we discuss the sample, and construct several tests to assure ourselves that sample attrition is not driving our results.

As student outcomes, we use scores in their end-of-year exams, the same measure as that used by the University of Cambridge to assess the students (see Section 2 for more details). The exams are held at the end of the academic year, one for every course the student took that year. We do not have access to scores on individual papers (except in Economics). Instead, we use the average percentage of marks obtained by the student across all of their exams taken that year, standardized by subject, as our outcome of interest. Since exams take place once a year over three years of their degree, we have three performance observations per student, one in each year of their degree.[18] Our initial analysis focuses on first-year performance and is then extended to subsequent years to examine the longer term validity of our main conclusions.

As discussed in Section 2.1, we focus on two student characteristics: gender and socioeconomic status. Gender is observed in our dataset. As proxy for socioeconomic status among UK students

[17] Economics, Engineering and Maths require applicants to have a minimum level of mathematics, e.g. Mathematics A-level in the UK school system, for their application to be considered. The same is required to enrol in Physical Sciences courses. In contrast, there is no such requirement for Biological Sciences, Law or Medicine. (University of Cambridge 2022b).

[18] In our sample, there are some exceptions to the three year rule: Engineering degree is four years long, Mathematics has an option to proceed to a fourth 'bonus' year and Medical degree is six years, including clinical study. For these, we use the first three years of exam performance for comparability.

we use whether they attended a state-funded or a privately-funded school, as discussed in Section 2.

Each of these characteristics has two categories: for gender, we label them h = female and g = male, for school-type, UK applicants from state-funded schools are labelled h and applicants from privately-funded schools g. Hence, in the language of Section 3.1.3, h is always the protected characteristic.

4.1 Gender

The gender composition of our sample is summarized in Table 1. In the subjects we analyze, 36% are female, but with marked differences across subjects. In non-mathematically intensive subjects, Law, Medicine and Biological Sciences, the number of females is close to or over 50%.[19] However, in mathematically intensive subjects it is significantly lower, at 24%. This pattern of female under-representation in mathematically intensive subjects is in line with what has been widely documented before in different settings cf. Wang and Degol 2017.

Turning to the probability of getting an offer from Cambridge in Table 2, across all subjects there is virtually no gender gap, with 23.9% and 23.2% respectively for males and females.[20,21] In mathematically intensive subjects, where they are overrepresented, males are slightly less likely to get an offer (23% vs 25%, and the gap is significant at 5% level), whilst the opposite is true in non-mathematically intensive subjects where they are underrepresented (24% vs 21%, and the gap is significant at 1% level).[22] These differences are suggestive that the decisionmaker is interested in gender balance as modelled in Section 3.1, but alone cannot be used to draw conclusions about the existence of performance-diversity trade-off.

We also see in Table 2 that, once admitted to Cambridge, males outperform females in the first-year exams in all subjects. The average first-year gap is 0.3 standard deviations (in terms of raw marks, which are expressed as percentages, this corresponds to about 3.5 percentage points difference, whereas the average first-year exam score is around 65%). The differences are signif-

[19] Overall, Cambridge undergraduate are evenly split by gender (University of Cambridge, 2022c). This is not the case in our sample, which does not include many smaller humanities and social science subjects where females are often in the majority.

[20] As discussed in Section 2, not all of the offer holders arrive, mostly because they fail to fulfil the conditions of their offer. We provide some data on this and discuss the implications for our analysis in Appendix C.1.

[21] For historical reasons, admission data treat applicants to Biological Sciences and Physical Sciences as one group, despite different requirements and process. For the main sample used in the paper, which is the admitted students who finish their three year degree, we can separate the two groups using data on their third year exams. However, for a few additional exercises when we need a larger sample, viz. calculating offer probabilities and some tests in appendixes, these two subjects are not separately identified. So, they are excluded from these in cases when we look at MI and Non-MI groups separately (but are still included in total numbers).

[22] These gaps in offer probabilities persist, and are slightly larger when we control for high school qualifications (see Appendix D).

Table 1: Gender composition of subjects

	N	Percent female	SE
	(1)	(2)	(3)
All subjects	5888	35.8	0.6
MI subjects	3544	24.4	0.7
Economics	547	33.8	2.0
Engineering	957	24.2	1.4
Mathematics	818	15.0	1.3
Physical Sciences	1222	26.6	1.3
Non-MI subjects	2344	53.0	1.0
Law	645	56.3	2.0
Medicine	895	45.0	1.7
Biological Sciences	804	59.3	1.7

Note. *N*: number of admitted students. *Percent female*: Percentage of female students in total admitted. *SE*: Standard error of Percent female. *MI subjects*: Mathematically intensive subjects, with math prerequisites. *Non-MI subjects*: Non-mathematically intensive subjects, without math prerequisites.

icantly larger in mathematically intensive subjects than in the rest. Among the former group of subjects, those with a lower proportion of females have a bigger gap between average male and female performance. These gender differences in average performance persist when we control for high school qualifications (Appendix I).

By the end of the third year, average student performance improves in all subjects. Furthermore, gender differences shrink, to an average of 0.1 standard deviations (or 0.9 percentage points relative to the average third year exam score of 66.5%). The gap closes in Biological Sciences, and is reversed in Medicine. In the rest of the subjects, the gender gap remains statistically significant.

Although these gender differences in *average* performance are suggestive, they cannot be used to draw conclusions about the diversity-performance trade-offs because the differences in average performance are generally not informative about the differences in performance of *marginal* candidates.

Table 2: Gender, offer probabilities and performance

	N	Percent female	Offer probability		Female score - Male score	
			Females	Males	Year 1	Year 3
	(1)	(2)	(3)	(4)	(5)	(6)
All subjects	5888	35.8	23.2	23.9	-0.32	-0.11
SE All		0.6	0.4	0.3	0.03	0.03
MI subjects	3544	24.4	25.0	23.4	-0.44	-0.25
SE MI		0.7	0.6	0.4	0.04	0.04
Non-MI subjects	2344	53.0	20.7	23.9	-0.25	0.02
SE Non-MI		1.0	0.6	0.6	0.04	0.04

Note. N: number of admitted students. *Percent female*: Percentage of female students in total admitted. *Offer probability*: Percentage receiving offer for admission, by gender. *Female score - Male score*: The difference between the standardized average exam score achieved by females and that achieved by males, in standard deviations. All differences are significant at 1% level except for Non-MI subjects Year 3, which are insignificant at conventional levels. *MI subjects*: Mathematically intensive subjects, viz. Economics, Engineering, Mathematics and Physical Sciences except in Offer Probability column which excludes Physical Sciences (see above) *non-MI subjects*: Non-mathematically intensive subjects, viz. Biological Sciences, Medicine and Law, except Offer Probability column which excludes Biological Sciences. *SE All/MI/Non-MI:* standard errors.

Table 3: School-type composition and performance

	N	Percent state	Offer probability		State score - Private score	
			State	Private	Year 1	Year 3
	(1)	(2)	(3)	(4)	(5)	(6)
Mean	4233	64.1	27.3	36.2	-0.14	-0.11
Standard error		0.7	0.4	0.6	0.03	0.03

Note. UK students only. N: number of admitted UK students. *Percent state*: Percentage of students who went to state-funded schools in total admitted. *Offer probability*: Percentage receiving offer for admission, by category. *State score - Private score*: The difference between the standardized average exam score achieved by students who went to state-funded schools and that achieved by students who went to privately-funded schools, in standard deviations.

4.2 School-type

To capture socioeconomic background, we use the UK government's main metric, which is whether the student went to a state- as opposed to a privately-funded school in the UK (see Section 2.1). Table 3 shows that in our data, 64% of admitted students come from state-funded schools, which happens to be the government's official target for Cambridge during this period. At the same time, students from privately-funded schools are more likely to receive an offer, and, once admitted, they outperform their peers from state-funded schools, on average.

4.3 Pool

Recall that the key to our empirical approach is the two-tier admission process, where some students are admitted directly (first tier) and others are admitted after being put in the pool (second tier). The latter group stands at about 1,215 students, or 20% of our sample.

Table 4 shows that in the first year, students admitted directly significantly outperform those admitted from the pool, by 0.2 standard deviations on average in all subjects. These differences shrink by the time the students reach their third year, but remain statistically significant, for both groups of subjects (MI and non-MI). The fact that candidates taken from the pool are weaker confirms that the two-tier admission system is working as intended and validates our key identifying assumption.

Table 4: Two tiers of admission

	N	Taken from the pool (%)	Exam score differences Direct - Pool	
			Year 1	Year 3
	(1)	(2)	(3)	(4)
All subjects	5888	20.6	0.21	0.13
SE All		0.5	0.03	0.03
MI subjects	3544	18.5	0.22	0.10
SE MI		0.7	0.04	0.04
Non-MI subjects	2344	23.9	0.20	0.18
SE non-MI		0.9	0.05	0.05

Note. *N*: number of admitted students. *Direct - Pooled*: The difference between the standardized average exam score achieved by directly admitted students and that achieved by students taken from the pool, in standard deviations, in Year 1 and Year 3 as denoted by column headings. *MI subjects*: Mathematically intensive subjects, viz. Economics, Engineering, Mathematics and Physical Sciences; *non-MI subjects*: Non-mathematically intensive subjects, viz. Biological Sciences, Medicine and Law. *SE All/MI/Non-MI*: standard errors.

5 Results: Gender

We now present our empirical findings on gender-performance trade-off, separately for subject groups based on their mathematical intensity.

5.1 Identifying and measuring the trade-off: mean comparisons

As discussed in Section 3.2.2, we identify performance-diversity trade-off by comparing mean exam performance across the groups. These (standardized) means of first-year exam performance at Cambridge are presented in Table 5, where g are males and h are females. Namely, the difference between average performance of pooled males (column (1)) and directly admitted females (column (2)) gives us the lower bound of the difference in admission thresholds and hence β, the weight on diversity in the university's objective function (column (3)). Whilst the difference between average performance of directly admitted males (column (4)) and pooled females (column (5)) gives us the estimate of the upper bound of β (column (6)). The standard errors reported here are equal to the estimates of the standard deviation of performance in each group (e.g. 'Pooled males in non-MI subjects') divided by the square root of the number of observations in that group.

Recall that we use the most conservative rule to identify the trade-off, i.e. we conclude there is one if and only if both the lower and the upper bounds of β are significantly greater than zero.

Table 5: Identifying and measuring the trade-off: Gender

	Pooled males (g)	DA females (h)	Lower bound of β	DA males (g)	Pooled females (h)	Upper bound of β
	(1)	(2)	(3)	(4)	(5)	(6)
non-MI subjects						
Mean	-0.04	-0.07	0.02	0.17	-0.23	0.39
SE	0.07	0.03	0.08	0.03	0.05	0.06
MI subjects						
Mean	-0.06	-0.30	0.25	0.14	-0.45	0.59
SE	0.04	0.04	0.06	0.02	0.06	0.07

Note. Pooled males (g)/females (h): average standardized performance in year 1 exams of students admitted from the pool from group g(males)/h(females). *DA males (g)/females(h)*: average standardized performance in year 1 exams of students directly admitted from group g(males)/group h(females). *Lower bound of β*: lower bound of the weight on diversity in the university objective function = Pooled g - DA h. *Upper bound of β*: upper bound of the weight on diversity in the university objective function = DA g - Pooled h. *MI subjects*: Mathematically intensive subjects, viz. Economics, Engineering, Mathematics and Physical Sciences; *Non-MI subjects*: Non-mathematically intensive subjects, viz. Biological Sciences, Medicine and Law. *SE*: standard errors, equal to the estimate of the standard deviation of performance divided by the square root of the number of observations, for each group.

Our conclusions from Table 5 are as follows:

Non-mathematical subjects: We find no conclusive evidence of a trade-off between gender equality and performance for these subjects. Although the upper bound of cutoff differences is positive and significant (column (6)), the lower bound is not significantly different from zero (column 3). Women represent just over 50% in these subjects, so, roughly speaking, gender equality has been achieved without any conclusive evidence that a performance sacrifice was involved in doing so. Recall that when we compare *average* outcomes, females slightly underperform males in these subjects (Table 2), illustrating that identifying a trade-off using average rather than marginal comparisons can lead to wrong conclusions.

Mathematically intensive subjects: There is a trade-off between gender diversity and performance in these subjects, as both the lower and the upper bounds of the differences in cutoffs are significantly above zero (columns 3 and 6). Marginal performances are not equalized. Instead, marginal males outperform marginal females by somewhere between 0.25 and 0.59 of a standard deviation. The former gives the lower bound on β, the diversity preference parameter in the decisionmaker's objective, i.e. the decisionmaker is prepared to forgo 0.25 standard deviations of performance to admit an additional female student at the margin. This implies that the decision-maker places a weight of $\frac{\beta}{1+\beta} = 20\%$ on admitting more women and 80% on academic performance in their objective function.

Contrast this result with a naive comparison of offer probabilities: as they are very close (23% for males and 25% for females), one might reach a mistaken conclusion that the admission cutoffs are set at a similar level for the two groups, and so there is no diversity-performance trade-off (see also Appendix D).

Remark 2 *Note that in all cases, the upper bound of the cutoff differences is positive which, combined with the sign of the lower bound estimates, implies that the pool system is working as intended, with pooled candidates weaker on average than directly admitted ones across all characteristics.*

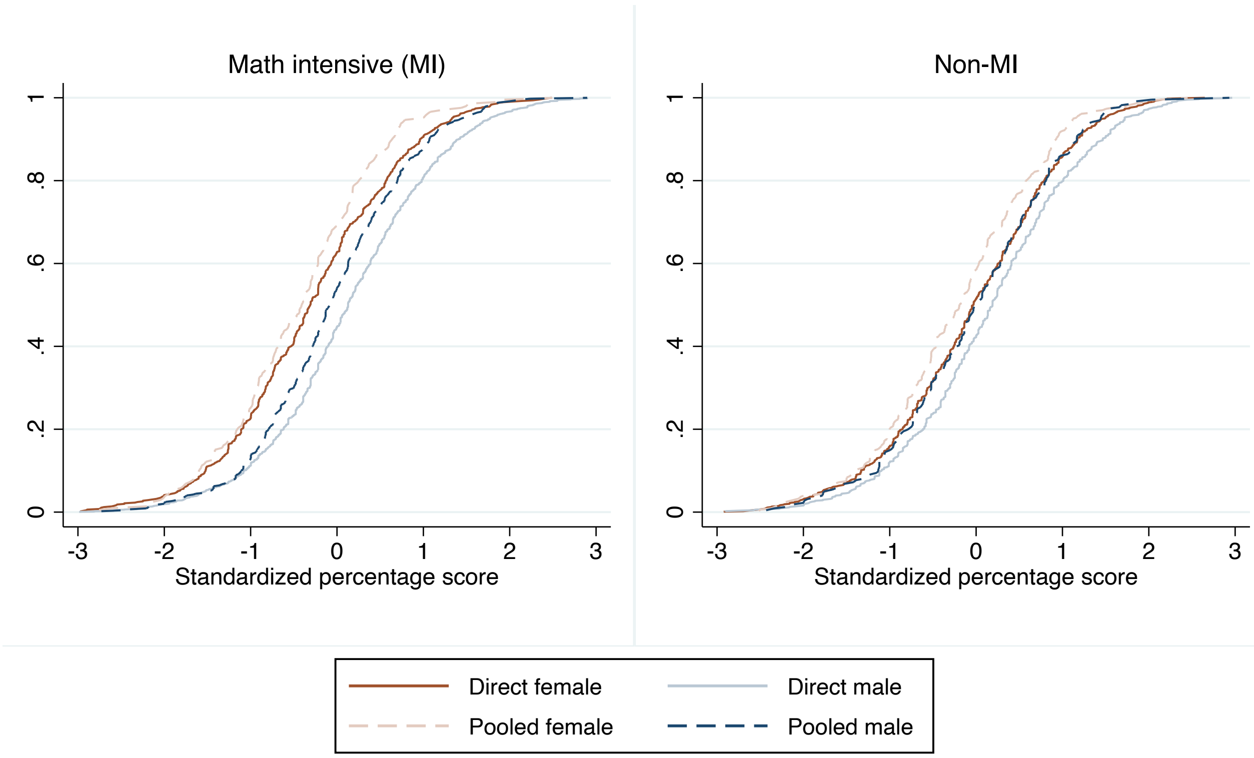
5.2 Distribution comparisons

In Figures 3, we plot the cumulative distribution function of first-year exam percentage scores for four subgroups of admitted students: pooled male, directly admitted male, pooled female and directly admitted female. The panel on the left is mathematically intensive subjects, the panel on the right is the rest.

In mathematically intensive subjects, we see clear evidence that the distribution of scores for directly admitted males first-order stochastically dominates the rest, followed by pooled males, directly admitted females and, finally, pooled females. The fact that pooled males have stochastically higher exam scores than directly admitted female entrants *throughout the distribution* suggests that $g_1 - h_1 > 0$, i.e. the de-facto admission cutoff is higher for males, seen in the light of equations (8) and (10) above. In contrast, the distribution of exam scores for pooled females is first-order stochastically dominated by the distributions of both pooled and directly admitted male scores. This confirms our finding that the university faces a gender-performance trade-off in these subjects. Figure 6 in Appendix F shows that the same pattern also holds in each individual subject in the mathematically intensive group i.e. Economics, Engineering, Mathematics and Physical Sciences. This also rules out the possibility that risk aversion can lead to the disparities we observe (cf. Section 3.3)

In contrast, in non-mathematically intensive subjects (Figure 3, right graph), the performance of pooled males and directly admitted females is similar, in line with our earlier finding that there is no gender-performance trade-off for these subjects. In Figure 7 in Appendix F we see this pattern in each of the individual subjects that comprise this group, viz. Biological Sciences, Law and Medicine.

Figure 3: First-year exam scores by pool status and gender, by subject group



Note. The graph shows the cumulative distribution function of first-year exam percentage scores for different sub-groups of students. The functions are plotted separately for MI and non-MI subjects

5.3 How large is the trade-off?

We can quantify the aggregate – as opposed to marginal – implications of our estimate of β through two back-of-the-envelope calculations. The first calculation asks if, using the *realized* post-entry performance data, we used the same cutoff for direct admissions from both groups, how much would average performance increase relative to the case where the cutoffs differ by the *magnitude of the lower bound* on $g_1 - h_1$ calculated above. The second calculation asks how many more h-types (and fewer g-types) are admitted because of the difference in cutoffs. In both calculations, it is implicitly assumed that potential applicants do not change their behavior in response to the change in admission strategies and that number and types of applicants stay fixed. In other words, the purpose of these exercises is to quantify the extent of misallocation, not to simulate a general equilibrium counterfactual. These calculations, derived in Appendix E, are 'back-of-the-envelope' in the sense that true post-entry performance is not observed by decisionmakers during admissions. What are observed are noisy measures of future performance, leading to some mis-classification. Also, this method assumes that the performance forgone is the same for every student that would not have been admitted using the ex-post measure, which may not be satisfied in reality.

These calculations reveal that relative to our estimate of the counterfactual with the same admission cutoff for all groups, the university directly admits around 10 additional girls into Mathematically Intensive subjects, or 6% of direct female intake (and, hence, does not directly admit the same number of males). In doing so, the university forgoes around 1.5 standard deviations in exam performance for each of these students.

5.4 Regression analysis and later years

We now use regression analysis to check robustness of our estimates on the lower bound of β, which is the critical part in identifying the performance-diversity trade-off (Table 6, Panel A). We also extend the analysis beyond first-year exams to later years (Table 6, Panel B).

In Table 6, our sample is the pooled males and directly admitted females. In panel A, we regress standardized first-year exam performance on a dummy indicating pooled males. Across all subjects (column (1)) this yields a positive, statistically significant coefficient on the dummy variable implying that pooled males score an average of 0.11 standard deviations higher than directly admitted female applicants.

Columns (3) and (4) confirm the previous finding that these results are being driven entirely by MI subjects. In these subjects, the first-year gap is 0.24 of standard deviation, significant at 1% level and very close to our mean comparison estimate in Table 5. This is the lower bound of the difference in admission cutoffs and confirms our earlier result that using even the most conservative measure, the university faces a gender-performance trade-off in these subjects.

With these estimates, we seek to address three key challenges. First, males could apply to subjects that are more selective in the first round of admissions than females, and so our mean comparison results may be due to differences in cutoffs across subjects rather than gender. We address this by including subject fixed effects (column (2) onwards). Second, decisions to admit and pool candidates are made by individual colleges. Figure 8 in Appendix G shows that colleges with a higher propensity to put male candidates into the pool tend to be better performing, on average, and so may be attracting stronger candidates. To address this we include application college fixed effects (column (3) onwards), and confirm that our result remains the same when we compare directly admitted females and pooled males who applied *to the same* college.

We also include offer college fixed effects in column (6) and the result remains significant, though it is smaller in magnitude and has a larger standard error. This is not surprising, since there are several large colleges that rarely take from the pool, and so including offer college fixed effects effectively reduces our sample by nearly 1/3. The reason to include these fixed effects would be if colleges that are more likely to take (male) students from the pool had, for example, better teaching or other resources and this could have led to higher exam performance. However,

Figure 9 in Appendix G shows that this concern is not empirically relevant: if anything, college 'quality' is inversely correlated with the propensity to take male applicants from the pool, and hence cannot explain the positive coefficients we obtain. For these reasons, we use the specification with subject fixed effects and application college fixed effects but without offer college fixed effects as our preferred one.

The students' first-year performance, whilst important, may not be the best measure of their overall performance at university. Therefore, we re-estimate our regressions using the student performance in their exams in the remaining two years of their university degree (Table 6, Panel B). Focussing on MI subjects only, first, we see that the statistically significant performance gap persists in the second and third years. Hence, using this longer run measure of performance, we continue to find existence of a diversity-performance trade-off. In magnitude, the gap shrinks, from 0.24 standard deviations in the first year to 0.15 standard deviations in the third year, which means that the implied lower bound of the diversity-performance trade-off is smaller when we look at performance in later years.[23]

Since our sample excludes students for whom exam data are missing and students who changed their course, we may ask whether our results are biased because of systematic attrition. In Appendix B, we show first that attriting pooled males and directly admitted females tend to be similar. Second, if we broaden our first year sample to include students that will drop out in future years, our findings remain unchanged. This gives us confidence that our results are not driven by systematic attrition.[24]

Finally, the performance of the pooled admits may not be the same as it would have been, had they attended the application college. But in order for this to affect our results, it has to be the case that male pooled candidates systematically end up in colleges that are better fit for them than the application college, while females do not. This seems unrealistic.

To summarize, our regression analysis controlling for a number of factors has confirmed our preliminary gender results reported in Table 6. There is no difference in lower bound of admission cutoffs for different genders in Law, Medicine and Biological Sciences, confirming that, using our conservative approach, we do not have sufficient evidence to identify diversity-performance trade-off in these disciplines. However, in mathematically intensive subjects (Mathematics, Engineering, Physical Sciences, and Economics), there is a significant difference in cutoffs, even using the lower bound estimates, thus implying that there is a gender equality-performance trade-off in these subjects. The lower bound of that trade-off signifies that at least 0.15-0.24 of a standard deviation in performance is given up at the margin to maintain the current proportion of female students at 24% in these disciplines.

[23]We also performed the same exercise for non-MI subjects and found that the estimates remain statistically insignificant in later years.

[24]We thank one of the anonymous referee for this suggestion.

Table 6: The gap between pooled males and directly admitted females

Panel A. First year

	All subjects		Non-MI	MI		
	(1)	(2)	(3)	(4)	(5)	(6)
Pooled male	0.11	0.17	0.03	0.27	0.24	0.15
	(0.05)	(0.06)	(0.07)	(0.06)	(0.06)	(0.08)
Observations	2235	2235	1129	1106	1106	1106
Subject FE		✓	✓	✓	✓	✓
Application college FE			✓		✓	✓
Offer college FE						✓

Panel B. Later years: Mathematically intesive subjects

	Second year	Third year
	(1)	(2)
Pooled male	0.16	0.15
	(0.06)	(0.06)
Observations	1106	1106
Subject FE	✓	✓
Application college FE	✓	✓

MI: Mathematically intensive subjects, viz. Economics, Engineering, Mathematics and Physical Sciences. *Non-MI subjects*: non-mathematically intensive subjects, viz. Biological Sciences, Medicine and Law. *Sample:* Pooled males and directly admitted females. Subjects are in column headers in Panel A, and MI in Panel B. *Dependent variable:* standardized exam score, 1st year (Panel A), 2nd and 3rd year (Panel B), as shown by column headers. *Pooled male:* a dummy = 1 if pooled male, = 0 if directly admitted female. *Estimation*: OLS. Standard errors clustered at application college level (columns 1 and 2, Panel A) and robust standard errors (columns 3-6, Panel A and all columns, Panel B) are in parentheses.

5.5 Robustness

We perform a few further robustness checks on our gender-gap results, which are reported in Table 7.

- Women-only colleges: Cambridge has three colleges that admit only women; these colleges were established later than most others, and tend to have lower financial resources. To check that our gender results are not driven by these, in addition to including application college fixed effects, we drop the women-only colleges from the sample (Table 7, column (2)).
- Other background controls: Another possible concern is false attribution, e.g. the gender gap stems from school-type if pooled females come mainly from state-funded schools whereas pooled males come mainly from privately-funded schools. To check this, we include dummies for school-type (UK state-funded, other UK, and non-UK) and the student's place of residence (EU, the UK, and other) as additional controls (Table 7, column (3)).[25]
- Year effects: We include fixed effects for year of application to ensure that the scores are comparable across years (Table 7, column (4)).

Table 7 shows that in MI subjects the gender gap is robust to all three issues discussed above.

Lastly, we consider an alternative way to identify the marginal candidates by focussing on a small number of pooled students who were eventually admitted by the same college that had placed them into the pool.[26] These students could be considered to be the marginal admits, since they had applied to and were not admitted by the pooling college in the first round, but were eventually admitted by them. Column (5) of Table 7 shows that, conditional on subject and application college fixed effects, marginal males score 0.23 standard deviations higher in first-year exams than marginal females. This result lends further support to our finding that the admission cutoff for females is lower than that for males. The number of such students is small (242), and so when we break this up by mathematical intensity of the subjects, the results are no longer statistically significant. Nevertheless, the point estimate is large and positive in both groups, and is about 1/3 higher in the MI subjects, consistent with our main results (columns (6) and (7)).

5.6 Understanding the gender gap

To help us further understand and interpret our main results in this section, we present additional findings on factors potentially related to the gender gap in mathematically intensive subjects. These

[25] The student's place of residence does not always coincide with the school location.

[26] 4/5 of applicants taken from the pool are taken by a different college from the one they had applied to, in line with the main rationale of the pool. The rest of the time, they are taken by the college they had applied to. This happens when a college puts a candidate in the pool but then takes them back, for example, if they could not find someone else they would rather take from the pool.

Table 7: Gender: robustness checks, year 1 performance

	All subjects					MI	non-MI
	Main	Mixed	Controls	Year FE	Marginal	Marginal	Marginal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Pooled male	0.24	0.24	0.24	0.25			
	(0.06)	(0.06)	(0.06)	(0.06)			
Marginal male					0.23	0.34	0.24
					(0.13)	(0.21)	(0.20)
Observations	1106	1052	1106	1106	242	122	120
Subject FE	✓	✓	✓	✓	✓	✓	✓
Application college FE	✓	✓	✓	✓	✓	✓	✓
Application year FE				✓			

Note. MI: Mathematically intensive subjects, viz. Economics, Engineering, Mathematics and Physical Sciences. *Non-MI subjects*: non-mathematically intensive subjects, viz. Biological Sciences, Medicine and Law. *Sample*: in columns except (1)-(4), the sample is pooled males and directly admitted females in MI subjects. In column (5), the sample consists of 'marginal' students in all subjects defined as students taken back from the pool by the college they applied to, further split by MI and non-MI subjects in columns (6) and (7). *Dependent variable*: the standardized score obtained in first-year exams. *Pooled male:* a dummy = 1 if pooled male, = 0 if directly admitted female. *Marginal male:* a dummy = 1 if marginal male, = 0 if marginal female. *Main:* main specification (table 6, panel A, column (5)). *Mixed:* only mixed gender colleges. *Controls:* additional controls for school-type and residence. *Year FE:* with fixed effects for year of application to Cambridge. *Estimation*: OLS. Robust standard errors are reported in parentheses.

factors can be grouped into two categories: those that are present at university and those in place prior to university admission.

5.6.1 At university

Peer Effects: In our sample, there are only 24% females in the mathematically intensive subjects, where we find the gender gap in performance. This is in contrast to the subjects where we do not identify a gender gap – in them, females are in a majority. Hence, it is possible that our results is due to peer effects: females are affected negatively by environments with fewer female classmates. This would challenge our interpretation of the main result as the difference in admission cutoffs. To address this we construct a test, exploiting the fact that a lot of teaching in Cambridge is in small peer groups, organized by college, and there is variation in gender composition across these. In column (1) of Table 8 we look at whether the gap in first-year performance is lower in subject-year-college combinations with a higher share of females. We do not find any evidence of such peer effects.

Compulsory and optional courses. Within each degree we analyze, there are some compulsory and some optional courses. We ask whether the gender gap is different in these two categories. Due to data limitations and complexity of some degrees, we focus on final year Economics exams, where there is a clear split into optional and compulsory papers. In column 2 of Table 8 we see that while pooled males still outperform directly admitted females in the compulsory courses (0.18 standard deviation, significant at 10% level), there are no performance differences in optional courses. This is consistent with the idea that once they are allowed to choose, females sort into the courses where they perform better. Since in most degrees there is a lot more choice in later compared to earlier years, this may partly explain our finding that the gap between pooled males and directly admitted females shrinks over time (in Table 6). This also suggests that the observed gender gap cannot be explained by some factor endemic to the university environment that disadvantages females across the board in their studies.

Table 8: University factors: Gender gap in mathematically intensive subjects

	Peer effects	Compulsory v optional courses
	(1)	(2)
Pooled male	0.26	0.04
	(0.12)	(0.11)
Share female	0.05	
	(0.27)	
Pooled male × Share female	-0.04	
	(0.37)	
Compulsory		-0.02
		(0.07)
Pooled male × Compulsory		0.18
		(0.10)
Observations	1106	1028
Subject FE	✓	
Application college FE	✓	✓

Note. *Sample*: Column 1: pooled males and directly admitted females in MI subjects. Column 2: pooled males and directly admitted females in Economics. *Dependent variable*: Column 1: the overall standardized score obtained in first-year exams in mathematically intensive subjects. Unit of observation: student. Column 2: the standardize score obtained in third-year exams for each course taken within the Economics degree. Unit of observation: student-course pair. *Pooled male:* a dummy = 1 if pooled male, = 0 if directly admitted female. *Share female*: for each student, we calculate the share of females in his/her subject-year-college combination (including that student). *Compulsory*: a dummy = 1 if the course is compulsory, 0 otherwise. *Estimation*: OLS. *Standard errors*: Robust in column (1); clustered at student level in column (2).

5.6.2 High school qualifications

We now ask whether the gender gap can be attributed to differences in high school qualifications observed at admissions. This is interesting in its own right, and also allows us to test implications of biased beliefs discussed in Section 3.4.

For comparability, we restrict our analysis to the students who have been through British-system schooling (80% of our sample). We focus on three high school qualification measures (explained below): summary grade obtained in general exams, grade in advanced mathematics exams, and number of advanced mathematics modules taken.

Typically, in British system, pupils apply to Cambridge when they are 17 years old, with the results from the two rounds of national school exams: General exams across a large number of subjects taken at 16 and advanced exams taken in specialist subjects at 17. More detailed information is in Appendix H.

Cambridge does not have cutoffs for these exam grades; instead applications are judged on a case-by-case basis. Both sets of exams are taken into account, but advanced exams are seen as more relevant because they are more recent, advanced, and specialist.

For general exams, we use the performance measure commonly used by top UK universities, including Cambridge, which is the number of subjects in which the applicant received the highest available grade (A*). This is what we use to measure general exam performance ('general exams' thereafter).

For advanced exams, admission decisionmakers focus on grades in individual subjects, particularly those most relevant to the applicant's future degree. Since mathematics is the one subject that is key to *all* mathematically intensive degrees in our study, we use the result of advanced maths exams ('math score') as our variable of interest. Our final variable is the number of advanced math modules, which reflects the fact that pupils can choose how many modules they study.[27]

Comparing high school qualifications by gender, males have lower scores than females in general exams, but higher math scores, and they also take more math modules, with the gender gaps in maths preparation larger for applicants to mathematically intensive subjects (Table 17 in Appendix H). These differences persists when we compare directly admitted females and pooled males.

We now ask whether the observed differences in high school qualifications are related to subsequent performance in Cambridge and whether they explain the gender gap in mathematically intensive subjects. To this end, we re-estimate our main regression (which compares the performance of pooled males and directly admitted females in mathematically intensive subjects), now controlling for high school qualifications and their interactions with gender, and report these results in Table 9. Since the estimates are conditional on admission, caution is required when interpreting them.

[27] Those taking more modules cover more material and cope with higher workload.

First, we verify that our main gender gap result holds qualitatively for the subset of students from the British school system (column (1)). In column (2), we add high school qualifications as regressors. Both general exams and maths scores (but not math modules) are positively and significantly correlated with performance, but the estimated performance gap between pooled males and directly admitted females does not change. Finally, in column (3) when we interact regressors with the type of student, some of the interactions are significant, but the gender performance gap, if anything, widens. We also use results in Table 9 to construct counterfactual predictions of Cambridge performance for *all* applicants to mathematically intensive subjects. Using these predictions and actual shares of admitted students, we show that the implied performance cutoff is lower for directly admitted females than for pooled males (Appendix I.4, Figure 10), echoing our main results found using actual performance.

To sum up, the results in Tables 9 show that, (i) both general exam results and math scores attained in school are correlated with performance in Cambridge; (ii) the relationship between high school qualifications and university performance differs by gender[28], and (iii) the performance gap between infra-marginal males and supra-marginal females does not appear to be explained by these differences. This is further confirmed by Oxaca decomposition of performance of pooled males and directly admitted females which we report in Appendix I.2. As we discuss in Section 3.4.2, one of the implications of this is that the gap we find between admission cutoffs for the two genders is unlikely to be driven by the university ignoring gender differences at admission.

Weights on qualifications

The coefficients in Table 9 (as well as in Table 18) suggest that advanced math scores are more strongly correlated with performance in mathematically intensive subjects than general exam scores. In Appendix I.3, we propose a rough way of testing whether the admission decisionmakers give appropriate weight to math scores when selecting candidates, and find some support for the idea that they tend to under-weight them relative to general exams. Taken together with the observation that females tend to perform better than males in general exams but worse in advanced math (Table 17), this lends some support to the notion that such under-weighting of math scores at admissions may contribute to the observed gender gaps in admission cutoffs (see also Section 3.4.3).

[28]Similar results are obtained if, instead of looking at infra-marginal males and supra-marginal females, we regress the performance of all directly admitted first-year students in MI subjects on gender dummy and pre-admission characteristics; see Table 18 in I.

Table 9: Qualifications: Gender gap in mathematically intensive subjects

	(1)	(2)	(3)
Pooled male	0.15	0.19	0.45
	(0.07)	(0.07)	(0.22)
Number of A* in General exams		0.09	0.14
		(0.04)	(0.05)
Maths score		0.33	0.42
		(0.04)	(0.05)
Maths modules		0.02	0.03
		(0.02)	(0.02)
Pooled male × Number of A* in General exams			-0.10
			(0.07)
Pooled male × Maths score			-0.23
			(0.09)
Pooled male × Maths modules			-0.04
			(0.03)
Observations	764	761	761
Subject FE	✓	✓	✓
Application college FE	✓	✓	✓

Note. *Sample:* Directly admitted females and pooled males enrolled in mathematically intensive subjects (Economics, Engineering, Mathematics and Physical Sciences) from British-system schools for whom both advanced school exam (AS) and general exam (GCSE) information is available. *Dependent variable:* the standardized score obtained in first-year Cambridge exams. *Pooled male:* a dummy variable that equals one for pooled males and zero for directly admitted females. *High school qualifications:* Number of A* in general exams – number of top marks (A*) achieved in general exams. Math score – percent math score obtained across all advanced (AS) math modules available at the time of application. *Estimation*: OLS. Robust standard errors are reported in parentheses.

6 Results: Socioeconomic background

Recall that, following the UK government definitions, we use state-funded school attendance as a proxy for lower socioeconomic background (see Section 2). First, we look at performance-diversity trade off by comparing mean exam performance as discussed in Section 3.2.2, where h is the protected characteristic, i.e. applicants from state-funded school.

Table 10, which for ease of interpretation has the same format as table 5, shows that we find no evidence of trade-off between socioeconomic diversity and performance. In both MI and non-MI subjects, although the upper bound of cutoff differences is positive and significant (column (6)), the lower bound is not significantly different from zero (column (3)). A zero lower bound does not rule out a positive parameter value, i.e. we may fail to detect an existing trade-off in this case, which is the cost of failure of point-identification. Note, however, that the share of students from state funded schools is 64%, which is the government target for Cambridge. Hence, we can say that the university has been able to achieve this target without any conclusive evidence that this involved a performance sacrifice, using our conservative measure.

Table 10: Identifying and measuring the trade-off: School-type

	Pooled private (g)	DA state (h)	Lower bound of β	DA private (g)	Pooled state (h)	Upper bound of β
	(1)	(2)	(3)	(4)	(5)	(6)
non-MI subjects						
Mean	-0.10	-0.02	-0.08	0.11	-0.20	0.32
SE	0.06	0.03	0.07	0.03	0.05	0.07
MI subjects						
Mean	-0.07	-0.07	-0.01	0.12	-0.31	0.44
SE	0.05	0.03	0.06	0.03	0.05	0.07

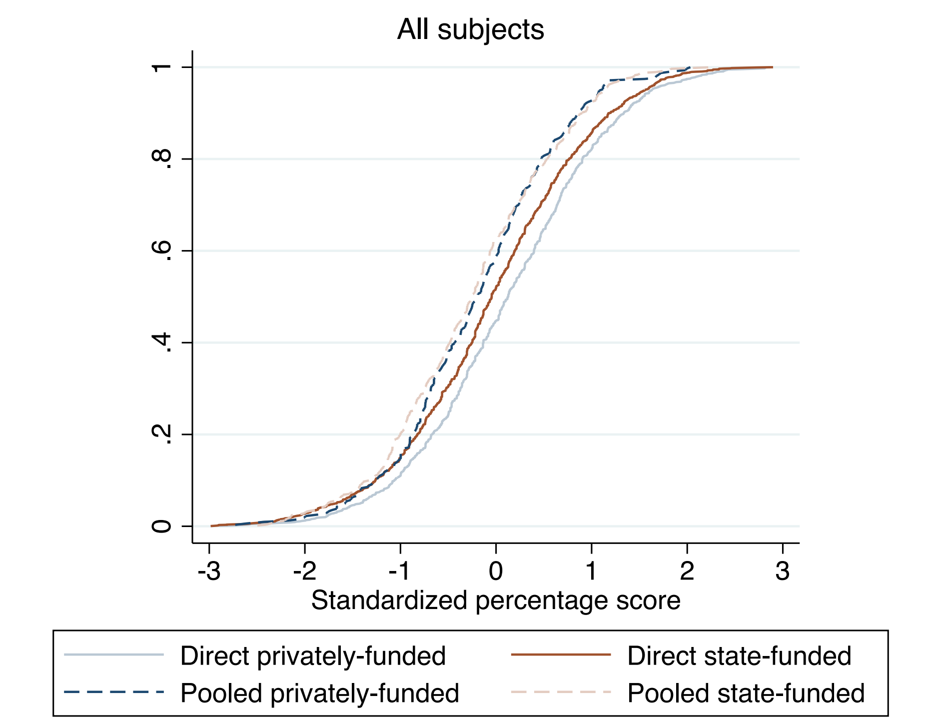
Note. Pooled private (g)/state (h): average standardized performance in year 1 exams of students admitted from the pool from group g(privately-funded schools)/h(state-funded schools). *DA private (g)/state(h)*: average standardized performance in year 1 exams of students directly admitted from the pool from group g(privately-funded schools)/h(state-funded schools). *Lower bound of β*: lower bound of the weight on diversity in the university objective function = Pooled g - DA h. *Upper bound of β*: upper bound of the weight on diversity in the university objective function = DA g - Pooled h. *SE*: standard error.

This conclusion is confirmed graphically in Figure 4: the CDF of first-year standardized exam scores for directly admitted distribution first order stochastically dominates the distribution of the pooled, within and across school-types. So, unlike in the case of gender in mathematically intensive subjects, the distribution of scores for directly admitted student with the protected characteristic is not FOSD by the pooled students who do not have the protected characteristic.

Robustness of this result is confirmed by regressions in Table 11, which control for subject and application college fixed effects. The results show that the mean differences are negative in all specifications across all years of exams, confirming our estimates in Table 5. Recall that negative differences do not provide evidence of a trade-off (sections 3.2 and 3.2.2). We also split the subjects into two groups by mathematical intensity, in the same way we have done for gender, and find no qualitative differences across them.

The fact that there is no positive difference in performance between super-marginal candidates from state-funded schools and marginal candidates from privately-funded schools implies that min $g_1 - h_1$, the lower bound of the diversity-performance trade-off, is not significantly different from zero. Hence, based on this conservative measure, there is no conclusive evidence of a trade-off between performance and socioeconomic status as captured by the school-type.

Figure 4: First-year exam scores by pool status and school-type



Note. The graph shows the cumulative distribution function of standardized first-year exam scores for different subgroups of students, by school-type, for all subjects combined.

Table 11: The school-type gap

	Year 1				Year 2	Year 3
	All subjects		Non-MI	MI	All subjects	
	(1)	(2)	(3)	(4)	(5)	(6)
Pooled privately-funded	-0.167	-0.207	-0.268	-0.127	-0.253	-0.136
	(0.053)	(0.053)	(0.068)	(0.081)	(0.058)	(0.053)
Observations	2459	2459	1431	1028	2459	2459
Subject FE		✓	✓	✓	✓	✓
Application college FE		✓	✓	✓	✓	✓

Note. *Sample*: directly admitted students from state-funded schools and pooled students from privately-funded schools (UK only). *Dependent variable*: indicated in the column header, where *Year 1, 2,* and *3* stand for standardized scores obtained in first-, second-, and third-year exams, respectively. *MI:* Mathematically intensive subjects, viz. Economics, Engineering, Mathematics and Physical Sciences. *Non-MI subjects*: non-mathematically intensive subjects, viz. Biological Sciences, Medicine and Law. *Pooled privately-funded:* a dummy = 1 for pooled candidates from privately-funded schools and 0 for directly admitted candidates from state-funded schools. *Estimation*: OLS. Standard errors clustered at the college of application level (columns (1) – (2)) and robust standard errors (columns (3) – (6)) are reported in parentheses.

7 Conclusion

Admission practices at elite universities face the dual objectives of maintaining high academic standards while admitting sufficiently many students from under-represented demographic groups. To make informed policy decisions it is important to understand whether there exist trade-offs between these objectives, and if so, what is their realized magnitude.

In this paper, we investigate this trade-off through a simple, intuitive model of admissions capturing the balance between diversity and post-entry student outcomes. We show that at the optimum in this model, the decisionmaker's implicit preference for specific demographics vis-a-vis expected future student outcome equals the difference in outcome between marginal admits across groups. This result is nonparametric in that it does not require assuming any specific functional form for distribution of unobservables, e.g. that potential outcomes are normally distributed or have right tails of a specific shape.

A challenge in implementing the above idea empirically is that it is usually not possible to determine who the marginal entrants are, when admission decisions are based on many indicators, some of which are unobserved by researchers, leading to well-known 'infra-marginality' problems. We develop a novel method to address this issue by exploiting the fact that many students enter elite universities, including Cambridge from where our data come, through a waitlist or second round clearing. We show that using inter-group outcome difference between students admitted directly versus through clearing, we can construct lower and upper bounds on the unobserved difference in the implicit admission cutoffs, i.e. the outcome of the marginal entrants. These bounds (a) do not require information on rejected candidates and (b) remain valid even if some applicant characteristics viewed by decisionmakers are unobserved by the researcher. This in turn enables us to back out the implied relative weight on demographics vis-a-vis future outcomes in the admission decisions. Given that a large majority of institutions in the world use waitlists and/or clearing to fill all their positions, this approach is likely to have wider applicability beyond the specific context studied here.

Finally, we apply our method to data from the University of Cambridge. We concentrate on two key applicant characteristics that decisionmakers typically focus on: gender and socioeconomic background as proxied by the type of school attended by the applicant. We use academic performance in Cambridge exams as the outcome of interest.

We find strong evidence that in mathematically intensive subjects where female enrolment is relatively low, there exists a significant trade-off between gender equality and performance. Waitlisted men outperform directly admitted females in exams by 0.25 standard deviations, which forms the lower bound of the future exam performance that the university gives up to prevent the gender ratio from deteriorating further in these subjects. For a standard normal benchmark, this suggests that between $10 - 15\%$ of admitted girls got in due to a lower cutoff. Further, this implies that the university places the relative weight of at least 20% on gender diversity in these subjects and of at most 80% on academic performance. Our measured performance gap is resilient to a large variety of robustness checks, persists throughout the length of the degree, especially in the compulsory core papers, and indicates a genuine underlying regularity.

On the other hand, we detect no gender-performance trade-off in non-mathematical but equally competitive subjects of Law, Medicine and Biological Sciences, where the gender ratio among admits is close to 50-50.

We also investigate the same trade-off for the case of socioeconomic background, as captured by the key variable on which the UK government has long based its admission guidelines, viz. the type of school attended by the candidate. We do not find strong evidence of trade-off between diversity and performance in this case. One important insight from these contrasting findings for gender and socioeconomic background is that the presence and magnitude of the trade-off are context-dependent.

Our results deliver a two-fold policy implication. First, to evaluate any policy intervention at the status-quo, the policy maker must have a measure of the relevant trade-offs. This can be challenging, and in this paper, we develop an approach for doing so. As we discuss in the introduction, all too often such trade offs are not made explicit, potentially leading to confused and inconsistent decision making. Our findings also stand in some contrast to the current policy debate in the UK, which focuses almost exclusively on the merits of increasing intake from state-schools, with relatively little attention to gender performance gap and recruitment into STEM fields. Our results suggest that in the latter area, the university faces much larger trade-offs, and so it should command more, not less, policy discussion.

This brings us to the second policy issue, which is how to address the gender performance gap in mathematically intensive subjects. This requires a good understanding of why the gap arises, a question that we cannot answer definitively with our data. Nevertheless, our evidence points to a gap in mathematical preparation existing already before entry into the university. Although females outperform males in general school exams, when it comes to mathematics preparation at a more advanced level, critical for mathematically intensive degrees at university, they lag somewhat behind their male counterparts.

Based on the above, we conjecture that in the short-run, boosting the mathematical preparation for female students admitted to STEM fields and Economics, say through a preparatory course, can potentially improve their subsequent performance and reduce the existing gender gap. Early childhood intervention encouraging bright girls to pursue mathematical tracks (cf. Heckman and Krueger 2005, Heckman 2006, Ellis et al. 2016, Wang and Degol 2017) would be a natural longer-term goal for increasing STEM opportunities and participation for women. Indeed, our results are consistent with the suggestion that girls are discouraged from taking as much maths as males in school and from applying to mathematically intensive subjects, with the academically stronger females going instead into disciplines like Law and Medicine (see Tonin and Wahba 2014 and Crawford et al. 2018). Indeed, it is conceivable that currently holding female applicants to a lower entry standard at admission, albeit at the expense of excluding more capable males, is a way to encourage better female students to apply to these subjects in the future, increasing overall efficiency in the long-run.

On the methodological end, our contributions are two-fold. Firstly, we develop a theoretically grounded empirical model of admissions that can be used to quantify precisely how the university trades off future outcomes against diversity objectives, and hence the relative weights it puts on the two. Secondly, we outline a novel method of set-identifying inter-group performance differences between marginal admits – an ingredient of our trade-off calculation – by using candidates admitted via waitlist, a common feature of college-admission around the world. These methods serve to mitigate two common complications in empirical assessment of selection procedures, viz. the unobservability of all applicant characteristics affecting selection decisions, and the lack of suitable counterfactual outcomes for rejected applicants. As such, our methods can be applied to other outcomes of interest, such as graduation rates, subsequent earnings, alumni donations etc. provided data on such outcomes are readily available.

References

[1] Albaek, K., 2017. Optimal admission to higher education. Education Economics, 25(1), pp.60-83.

[2] Alesina, A. and E. La Ferrara, 2014. A Test of racial bias in capital sentencing, American Economic Review, 104, pp.3397–3433.

[3] Altonji, J.G. and Blank, R.M., 1999. Race and gender in the labor market. Handbook of labor economics, 3, pp.3143-3259.

[4] Anwar, S. and H. Fang, 2006. An Alternative test of racial prejudice in motor vehicle searches: Theory and evidence, American Economic Review, 96, pp.127–151.

[5] Arcidiacono, P., Lovenheim, M. and Zhu, M., 2015. Affirmative action in undergraduate education. Annual Review of Economics, 7(1), pp.487-518.

[6] Arcidiacono, P. and Lovenheim, M., 2016. Affirmative action and the quality-fit trade-off. Journal of Economic Literature, 54(1), pp.3-51.

[7] Arcidiacono, P., Aucejo, E.M. and Hotz, V.J., 2016. University differences in the graduation of minorities in STEM fields: Evidence from California. American Economic Review, 106(3), pp.525-62.

[8] Arcidiacono, P., Kinsler, J. and Ransom, T., 2023. What the Students for Fair Admissions Cases Reveal About Racial Preferences. Journal of Political Economy Microeconomics, 1(4), pp.000-000.

[9] Arrow, K. J. 1973. The Theory of Discrimination., In Orley Ashenfelter and Albert Rees, Discrimination in Labor Markets. Princeton, N.J.: Princeton University Press, 1973, pp. 3–33.

[10] Autor, D.H. and Scarborough, D., 2008. Does job testing harm minority workers? Evidence from retail establishments. The Quarterly Journal of Economics, 123(1), pp.219-277.

[11] Ayres, I., 2002. Outcome Tests of Racial Disparities in Police Practices, Justice Research and Policy, 4, pp.131–142.

[12] Ayres, I. and Brooks, R., 2004. Does affirmative action reduce the number of black lawyers. Stanford Law Review, 57, p.1807.

[13] Becker, G.S., 1957. The economics of discrimination. University of Chicago Press.

[14] Bertrand, M., Hanna, R. and Mullainathan, S., 2010. Affirmative action in education: Evidence from engineering college admission in India. Journal of Public Economics, 94(1), pp.16-20.

[15] Bhattacharya, D. and Dupas, P., 2012. Inferring welfare maximizing treatment assignment under budget constraints. Journal of Econometrics, 167(1), pp.168-196.

[16] Bhattacharya, D., Kanaya, S. and Stevens, M., 2017. Are university admissions academically fair? Review of Economics and Statistics, 99(3), pp.449-464.

[17] Black, S.E., Denning, J.T. and Rothstein, J., 2023. Winners and losers? The effect of gaining and losing access to selective colleges on education and labor market outcomes. American Economic Journal: Applied Economics, 15 (1), pp.26-67.

[18] Bleemer, Z., 2021. Top Percent Policies and the Return to Postsecondary Selectivity, mimeo.

[19] Bleemer, Z., 2022. Affirmative action, mismatch, and economic mobility after California's Proposition 209. The Quarterly Journal of Economics, 137(1), pp.115-160.

[20] Bohren, J.A., Haggag, K., Imas, A. and Pope, D.G., 2019. Inaccurate statistical discrimination: An identification problem (No. w25935). National Bureau of Economic Research.

[21] Boliver, V., 2013. How fair is access to more prestigious UK universities?. The British Journal of Sociology, 64(2), pp.344-364.

[22] Bostwick, V.K. and Weinberg, B.A., 2022. Nevertheless she persisted? Gender peer effects in doctoral STEM programs. Journal of Labor Economics, 40(2), pp.397-436.

[23] Canay, I.A., Mogstad, M. and Mountjoy, J., 2020. On the use of outcome tests for detecting bias in decision making (No. w27802). National Bureau of Economic Research.

[24] Crawford, C., Davies, N. M. and Smith, S. 2018. Why do so few women study economics? Evidence from England. Royal Economic Society working paper.

[25] Cummings, Jack, 2015. Girls 'lack confidence' for Oxbridge. The Tab. https://thetab.com/2015/09/25/girls-lack-the-confidence-for-oxbridge-54920

[26] Ellis, J., Fosdick, B.K. and Rasmussen, C., 2016. Women 1.5 times more likely to leave STEM pipeline after calculus compared to men: Lack of mathematical confidence a potential culprit. PloS one, 11(7).

[27] Fryer Jr, R.G. and Loury, G.C., 2005. Affirmative action and its mythology. Journal of Economic Perspectives, 19(3), pp.147-162.

[28] Heckman, J.J., 2006. Skill formation and the economics of investing in disadvantaged children. Science, 312(5782), pp.1900-1902.

[29] Heckman, J.J. and Krueger, A.B., 2005. Inequality in America: What role for human capital policies?. MIT Press Books.

[30] Hoffman, M., Kahn L. B., and D. Lee, 2018. Discretion in Hiring. Quarterly Journal of Economics, 133(2), pp.765-800.

[31] Hull, P., 2021. What marginal outcome tests can tell us about racially biased decision-making (No. w28503). National Bureau of Economic Research.

[32] Ingrey, J., 2021. The maths faculty is failing its female students – literally. Varsity, October 29, 2021. https://www.varsity.co.uk/opinion/22298. Accessed 09/04/2023.

[33] Keith, S.N., Bell, R.M., Swanson, A.G. and Williams, A.P., 1985. Effects of affirmative action in medical schools: a study of the class of 1975. New England Journal of Medicine, 313(24), pp.1519-1525.

[34] Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J. and Mullainathan, S., 2018a. Human decisions and machine predictions. The Quarterly Journal of Economics, 133(1), pp.237-293.

[35] Kleinberg, J., Ludwig, J., Mullainathan, S., and A. Rambachan 2018b. Algorithmic Fairness. AEA Papers and Proceedings, 108, pp.22-27.

[36] Kleinberg, J., Mullainathan, S., and M. Raghavan 2017. Proceedings of the 8th Conference on Innovations in Theoretical Computer Science (ITCS)

[37] Knowles, J., Persico, N. and Todd, P., 2001. Racial bias in motor vehicle searches: Theory and evidence. Journal of Political Economy, 109(1), pp.203-229.

[38] Ladd, H.F., 1998. Evidence on discrimination in mortgage lending. Journal of Economic Perspectives, 12(2), pp.41-62.

[39] Sander, R.H., 2004. A systemic analysis of affirmative action in American law schools. Stanford Law Review, 57, p.367.

[40] Sibieta, Luke 2021. The growing gap between state school and private school spending. Institute for Fiscal Studies Observation, 08 October 2021 https://ifs.org.uk/publications/15672

[41] Stoye, J. 2021. A Simple, Short, but Never-Empty Confidence Interval for Partially Identified Parameters. https://arxiv.org/abs/2010.10484

[42] Sutton Trust 2018: https://www.suttontrust.com/wp-content/uploads/2019/12/AccesstoAdvantage-2018.pdf

[43] Tonin, M. and J. Wahba 2014. The Sources of the Gender Gap in Economics Enrolment. IZA Discussion Paper No. 8414.

[44] United States Supreme Court 2023: www.theguardian.com/law/2023/jun/29/us-supreme-court-affirmative-action-harvard-unc-ruling

[45] University of Cambridge 2019. Access Agreement with the Office for Fair Access (OFFA) 2018-19.

[46] University of Cambridge 2020. Access and Participation plan.

[47] University of Cambridge 2022a. Course list. https://www.undergraduate.study.cam.ac.uk/courses/

[48] University of Cambridge 2022b. Natural Sciences. Course outline. https://www.undergraduate.study.cam.ac.uk/courses/natural-sciences#course-outline

[49] University of Cambridge 2022c. Undergraduate Admissions Statistics, 2021 cycle

[50] University of Cambridge 2023. Admissions Policy

https://www.undergraduate.study.cam.ac.uk/applying/decisions/admissions-policy

[51] Wang, M., and J. L. Degol 2017. Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. Educational Psychology Review 29(1), pp. 119-140.

[52] Zimdars, A., 2010. Fairness and undergraduate admission: a qualitative exploration of admissions choices at the University of Oxford. Oxford Review of Education, 36(3), pp.307-323.

Appendices

A Proofs of Claims 1 and 2

Proof of Claim 1. Given any set $\mathcal{X}_h$, choose h_1 as the $\Pr(X \notin \mathcal{X}_h | G = h)$th quantile of $E(Y|X, G = h)$, i.e.

$$\Pr\{E(Y|X, G = h) \geq h_1\} = \Pr(X \in \mathcal{X}_h | G = h),$$

implying

$$\begin{aligned}
& N_h \Pr(X \in \mathcal{X}_h, E(Y|X, G = h) < h_1) + N_h \Pr(X \in \mathcal{X}_h, E(Y|X, G = h) \geq h_1, G = h) \\
= \; & N_h \Pr\{E(Y|X, G = h) \geq h_1, X \in \mathcal{X}_h | G = h\} \\
& + N_h \Pr\{E(Y|X, G = h) \geq h_1, X \notin \mathcal{X}_h | G = h\},
\end{aligned}$$

i.e.

$$N_h \Pr(X \in \mathcal{X}_h, E(Y|X, G = h) < h_1 | G = h) = N_h \Pr\{E(Y|X, G = h) \geq h_1, X \notin \mathcal{X}_h | G = h\}. \tag{12}$$

Therefore, the difference in expected outcomes between the two alternative criteria for admissions, viz. $\{X \in \mathcal{X}_h\}$ and $\{E(Y|X, G = h) \geq h_1\}$ is given by

$$\begin{aligned}
& N_h E(Y \times 1\{X \in \mathcal{X}_h\} | G = h) + \beta N_h \Pr(X \in \mathcal{X}_h | G = h) \\
& - N_h E(Y 1\{E(Y|X, G = h) \geq h_1\} | G = h) - \beta N_h \Pr(E(Y|X, G = h) \geq h_1) \\
= \; & N_h E((Y + \beta) \times 1\{X \in \mathcal{X}_h\} | G = h) \\
& - N_h E((Y + \beta) \times 1\{E(Y|X, G = h) \geq h_1\} | G = h) \\
= \; & N_h E((Y + \beta) \times 1\{X \in \mathcal{X}_h, E(Y|X, G = h) < h_1\} | G = h) \\
& - N_h E((Y + \beta) \times 1\{X \notin \mathcal{X}_h, E(Y|X, G = h) \geq h_1\} | G = h) \\
< \; & N_h (h_1 + \beta) \times \Pr\{X \in \mathcal{X}_h, E(Y|X, G = h) < h_1 | G = h\} \\
& - N_h (h_1 + \beta) \times \Pr\{X \notin \mathcal{X}_h, E(Y|X, G = h) \geq h_1 | G = h\} \\
= \; & N_h (h_1 + \beta) \times \underbrace{\begin{bmatrix} \Pr\{X \in \mathcal{X}_h, E(Y|X, G = h) < h_1 | G = h\} \\ - \Pr\{X \notin \mathcal{X}_h, E(Y|X, G = h) \geq h_1 | G = h\} \end{bmatrix}}_{=0, \text{ by } (12)} \\
= \; & 0.
\end{aligned}$$

This implies that the criterion $\{X \in \mathcal{X}_h\}$ yields a lower expected outcome than the criterion $\{E(Y|X, G = h) \geq h_1\}$.

For group g, one can repeat the above argument after replacing h by g and setting $\beta = 0$, to conclude that picking g_1 that solves

$$\Pr\{E(Y|X, G = g) \geq g_1\} = \Pr(X \in \mathcal{X}_g | G = g),$$

for any other subset $\mathcal{X}_g$ of the support of X_g will imply that

$$N_g E\left(Y \times 1\left\{X \in \mathcal{X}_g\right\}|G=g\right) < N_g E\left(Y \times 1\left\{E\left(Y|X_g, G=g\right) \geq g_1\right\}|G=g\right).$$

Therefore, the optimal solutions $\mathcal{X}_g, \mathcal{X}_h$ to the problem (16) must be of the form $\mathcal{X}_g = \{E\left(Y|X, G=g\right) \geq g_1\}$ and $\mathcal{X}_h = \{E\left(Y|X, G=h\right) \geq h_1\}$. ■

Proof of Claim 2. Let $r = 1 - F_g\left(g_1\right)$ and $s = 1 - F_h\left(h_1\right)$. Since $F_g\left(\cdot\right)$, $F_h\left(\cdot\right)$ are strictly increasing and continuous, maximizing with respect to g_1, h_1 is equivalent to maximizing w.r.t. r, s. Therefore, (3) is equivalent to

$$\max_{r,s\in[0,1]}\left[N_g \int_{F_g^{-1}(1-r)}^{\infty} a f_g\left(a\right) da + N_h \int_{F_h^{-1}(1-s)}^{\infty} a f_h\left(a\right) da + \beta N_h s\right]$$
$$s.t.$$
$$N_g r + N_h s = M.$$

Replacing $r = \frac{M-N_h s}{N_g}$, the objective becomes

$$\left[N_g \int_{F_g^{-1}\left(1-\frac{M-N_h s}{N_g}\right)}^{\infty} a f_g\left(a\right) da + N_h \int_{F_h^{-1}(1-s)}^{\infty} a f_h\left(a\right) da + \beta N_h s\right] \tag{13}$$

F.O.C. for maximum

$$-N_h \frac{F_g^{-1}\left(1-\frac{M-N_h s}{N_g}\right) f_g\left(F_g^{-1}\left(1-\frac{M-N_h s}{N_g}\right)\right)}{f_g\left(F_g^{-1}\left(1-\frac{M-N_h s}{N_g}\right)\right)} + N_h \frac{F_h^{-1}\left(1-s\right) f_h\left(F_h^{-1}\left(1-s\right)\right)}{f_h\left(F_h^{-1}\left(1-s\right)\right)} + \beta N_h = 0$$

i.e.

$$\beta = F_g^{-1}\left(1-\frac{M-N_h s}{N_g}\right) - F_h^{-1}\left(1-s\right). \tag{14}$$

As for the second order condition, note that

$$\begin{aligned} & \frac{\partial}{\partial s}\left\{-N_h F_g^{-1}\left(1-\frac{M-N_h s}{N_g}\right) + N_h F_h^{-1}\left(1-s\right) + \beta N_h\right\} \\ = \quad & -\frac{N_h^2}{N_g}\frac{1}{f_g\left(F_g^{-1}\left(1-\frac{M-N_h s}{N_g}\right)\right)} - \frac{N_h}{f_h\left(F_h^{-1}\left(1-s\right)\right)} \\ < \quad & 0 \text{ for all } s. \end{aligned}$$

So, the objective function (13) is a strictly concave function of s, and thus the first order condition (14) yields a unique maximum at an interior point. Now, note that $F_h^{-1}\left(1-s\right) = h_1$ and $F_g^{-1}\left(1-\frac{M-N_h s}{N_g}\right) = F_g^{-1}\left(1-r\right) = g_1$, the admission cutoffs for groups h and g respectively. Therefore, (14) reduces to

$$\beta = g_1 - h_1. \tag{15}$$

■

B Our model as Extended Roy Model

Canay et al. 2022 set up the problem of outcome based tests of decision-making (illustrated via judicial verdict on whether to grant bail) through Extended Roy Model.

Using their notation in our context, denote $D_i = 1$ as the decision to admit applicant i and $D_i = 0$ the decision to reject, the groups g, h to be values assumed by the commonly observed characteristic R_i, the potential outcomes Y_{1i} equal to i's potential academic performance upon entering the university, and $Y_{0i} \equiv 0$, reflecting that the university is assumed not to care about the applicant's potential performance (in another university) if he/she is not admitted. Like Canay et al. 2022, we maintain the assumption that the university's decision does not affect the values of potential outcomes. In our context, this rules out for example that one group exerts additional effort upon entry if they realize that they had faced higher or lower standards of admission. The 'cost' C associated with the decision of $D_i = 0$ is 0 if $R_i = g$ and equals $-\beta$ if $R_i = h$. That is, the university enjoys an additional (psychic) gain of β, measured in outcome units, when accepting an h-type applicant.

Additionally, there is a constraint on the total number of admits, forced by the capacity constraint, which is not explicitly modelled in Canay et al. As a result, the university's decision process has to account for not just whether the net benefit from its decision is positive as in Canay et al., but also for whom are the net benefits the highest, leading to the group specific thresholds g_1 and h_1. Subject to that modification, equation (15) above is somewhat analogous to equation (16) in Theorem 4.2 of Canay et al. and essentially reproduces the Beckerian insight that the difference in outcomes of marginal entrants from different groups reveals the difference in standards to which they were held; a larger outcome for group g marginals implies that they are held to a higher standard.

C Sample and attrition

Our sample consists of all students who entered Cambridge in 2013-2016 and subsequently sat exams at the end of the three year degree in seven large subjects (Economics, Engineering, Mathematics, Biological Sciences, Law, Medicine and Physical Sciences). This is 5,888 students (see also Section 4).

Figure 5 summarizes Cambridge admission process, whilst Table 12 shows how the sample above is arrived at. In 2013-2016, Cambridge received roughly 38 thousand applications, of whom around 9 thousand were given admission offers. Subsequently, around 7 thousand of them enrolled, with the remaining 2 thousand either missing their conditional offer or choosing not to come to Cambridge. Of the 7 thousand enrolled, around one thousand students did not reach the end of their degree, resulting in the final sample of just under 6 thousand students.

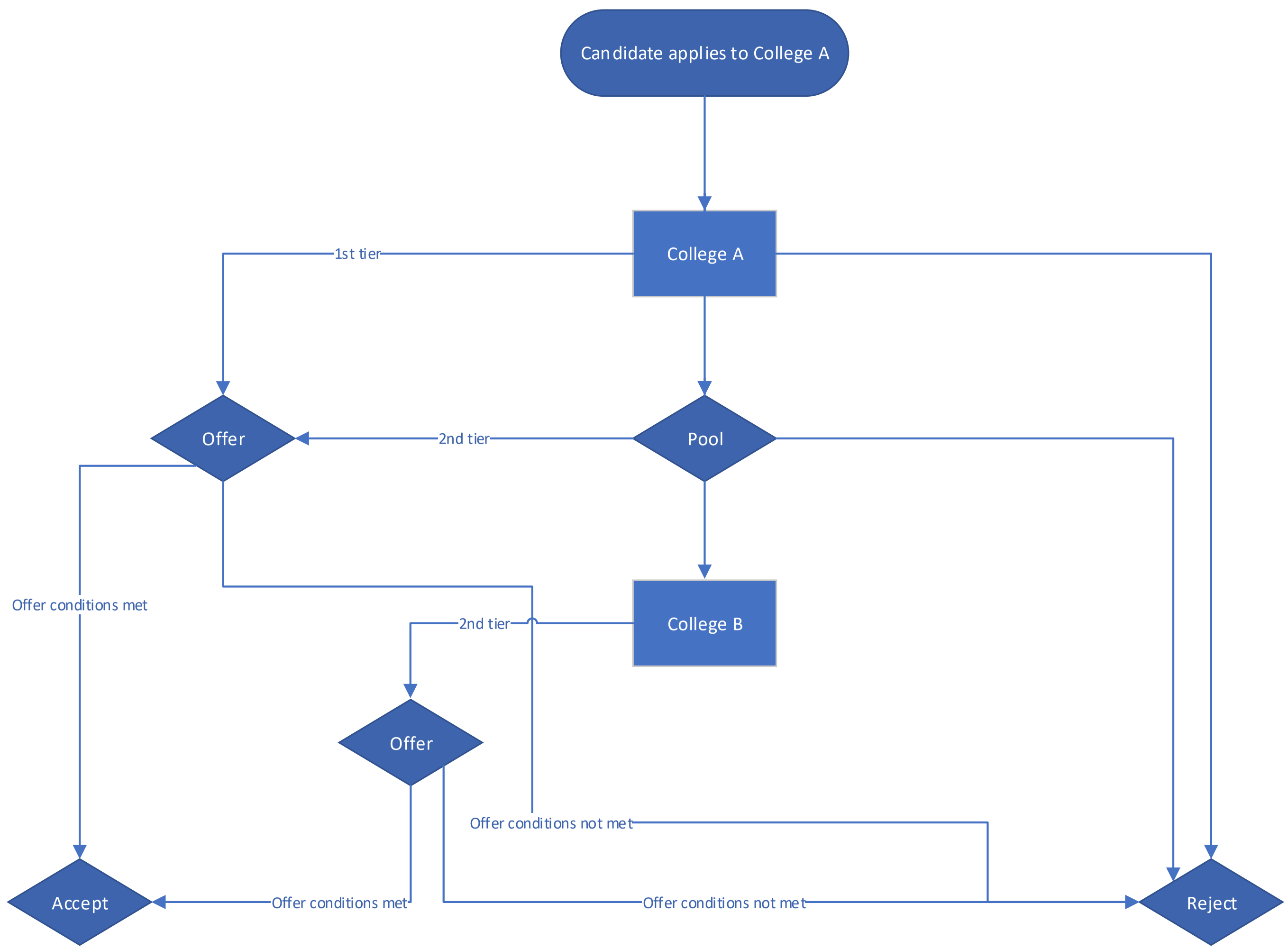
This attrition between the offers made and our sample, though unavoidable in our outcome-based analysis, may raise concerns. At the same time, in order for attrition to be a threat to our results, the probability of attrition has to be correlated with performance and group membership in very particular ways. In the subsequent Sections, we discuss the sources of attrition in our context, and present evidence that such correlations are unlikely. We do this in two steps: Section C.1 discusses pre-enrolment attrition (from 9 to 7 thousand, roughly), and Section C.2, post enrolment attrition (from 7 to 6 thousand, roughly).

Table 12: Sample

	N	%
Applied to Cambridge	38,199	
Recieved offers	9,028	24
Enrolled	7,089	79
Exams in Year 1	6,732	95
Exams in Year 3 (final sample)	5,888	87

Note. Exams in Year 1/Year 3 are the number of students who sat exams in the subject they enrolled in.
% column shows the N entry in this row as a percent of the N entry in the previous row (e.g. 87% in the last row means that 87% of all those who sat exams in the subject of their initial enrolment at the end of year 1 also did so at the end of year 3).

Figure 5: The two tiers of Cambridge admission process



C.1 Pre-enrolment attrition

Table 12 shows that just over 20% of applicants who receive an offer from Cambridge do not end up enrolling. This can happen for two reasons, an offer holder missing their conditional offer[29] or an offer holder choosing not to come. Although unfortunately the university does not collect systematic data on this, because of high admission requirements and elite status of the university, the former is a considerably greater concern to the university's admission officers than the latter. In particular, when it comes to domestic applicants (who account for over 2/3 of the student intake, and pay relatively low domestic fees), the operational assumption in Cambridge is that probability of them turning down a Cambridge offer is negligible. [30]

Potentially, pre-enrolment attrition could impact our analysis in two ways: first, through our model of the diversity-performance trade-off in Section 3.1 and second, through empirical identification of admission cutoffs in Section 3.2. We argue below that neither of these affect the main results of the paper.

To address the first concern, we can introduce arrival probabilities below 100% into our decisionmakers model. It can be shown that, even if they differ across the groups of interest, our main result, summarized in equation (15) is not affected (see below). Intuitively, arrival rates smaller than 100% will force the decisionmaker to take more students than the M places available. In doing so, the decisionmaker will have to lower the marginal expected performance of those admitted, i.e. lower g_1 and h_1. Nevertheless, on the margin the decisionmaker will continue to equate the marginal benefit from admitting from h group $(h_1 + \beta)$ to that of admitting from g group (g_1); if this condition does not hold, the decisionmaker could change the composition of admitted to increase his objective. To see this formally, let the probability of acceptance be P for a randomly chosen applicant (a special case is where the probabilities are identical across all type g and type h applicants, respectively, i.e. $P = p_g$ with probability 1 if $G = g$, and $P = p_h$ with probability 1 if $G = h$), then the optimization problem becomes:

$$\max_{\mathcal{X}_g,\mathcal{X}_h} \left[\begin{array}{c} N_g E\left(PY \times 1\left\{X \in \mathcal{X}_g\right\}|G = g\right) + N_h E\left(PY \times 1\left\{X \in \mathcal{X}_h\right\}|G = h\right) \\ +\beta N_h E\left\{P \times 1\left(X \in \mathcal{X}_h\right)|G = h\right\} \end{array} \right], \tag{16}$$

subject to

$$N_g E\left\{P \times 1\{X \in \mathcal{X}_g\}|G = g\right\} + N_h E\left\{P \times 1\left\{X \in \mathcal{X}_h\right\}|G = h\right\} = M. \tag{17}$$

[29] As we discuss in Section 2.3, Cambridge offers are made early – in January for the coming October. As a result, they are almost always conditional on obtaining a certain mark in school leaving exams (usually sat in the summer), and, in the special case of Mathematics, on an entrance exam.

[30] The one UK university that is a clear rival to Cambridge is Oxford. However, the rules governing UK applications do not allow students to apply to both Cambridge and Oxford in the same year.

The condition $\beta = g_1 - h_1$ will continue to hold in the transformed problem defined by (16) and (17), but g_1 and h_1 will be thresholds for $E[PY|X, G = g]$ and $E[PY|X, G = h]$, not $E[Y|X, G = g]$ and $E[Y|X, G = h]$. The lower bound on β will then be $E[PY|\text{Waitlist admitted } g] - E[PY|\text{Directly admitted } h]$.

The first expectation can be consistently estimated by the average performance of all g-type applicants offered admission from the waitlist, where the performance of those who did not accept is counted as 0. Analogously for the second term.

For the second concern, first note that pre-enrolment attrition that is due to offer holders not meeting their offer, which is the bigger concern in Cambridge, is not a threat to empirically identifying the difference in cutoffs for direct admission. This is because the offer holders who do not meet the conditions of their offer are *rejected* by the university.[31] Since the university is not admitting these students, their removal from our sample is fully justified for the estimation of admission cutoffs.

Finally, consider attrition of the smaller group of (predominantly international) students who choose not to enrol. This presents a threat to identification only if there are very particular correlations between such attrition, subsequent performance and group membership. Specifically, for such attrition to explain our main result – that the admission cutoff is lower for females in MI subjects – directly admitted female offer holders who choose not to enrol must be (a) stronger than pooled male offer holders choosing not to enrol and (b) this effect has to be large enough to change the sign of the difference in mean performance between directly admitted females and pooled males. Although it might be possible to construct scenarios when this would be the case, there are no obvious intuitive reasons to expect such patterns. Also, when we re-estimate our main regressions on the subsample of domestic students, for whom this type of attrition is very rare, our results are quantitatively smaller, but qualitatively the same. This gives us further confidence that they are not driven by such attrition.

[31] Although we do not model this explicitly, not meeting offer conditions can be thought of as the university updating the expectations of students' ability, A, such that $E(A|\text{conditions not met}) < g_2$ (or h_2 depending on the group), as defined in Section 3.2

Table 13: The gender gap in MI subjects, year 1 performance

	Full sample	Domestic students only
	(1)	(2)
Pooled male	0.24	0.14
	(0.06)	(0.07)
Observations	1106	730
Subject FE	✓	✓
Application college FE	✓	✓

Note. Each column reports the results from a different OLS regression. *Pooled male*: a dummy variable that equals one for pooled males and zero for directly admitted females. Sample: Column (1) All pooled males and directly admitted females enrolled in mathematically intensive subjects (same estimation as in Table 7, column (7)); Column (2) Domestic pooled males and directly admitted females enrolled in mathematically intensive subjects. *Dependent variable*: the standardized score obtained in exams, in first-year. *Estimation*: OLS. Robust standard errors are in parentheses.

C.2 Post-enrolment attrition

In the paper, to keep our sample stable, we restrict it only to the students for whom we have all three years of examination data in the subject they enrol in. Thus, over time we 'lose' students who drop out, change subject, or both (although recall that, unlike for example in the US, in the UK students decide on their University subject at application, and it is difficult to change it later, so this is a fairly rare phenomenon (see Section 2.3). As we see in Table 12, we start with 7,089 students who accept Cambridge's offer, and we subsequently lose 1,201 of them. Most of this attrition occurs in years 2 and 3, when we 'lose' 844.

It is, therefore, a legitimate question whether our findings are driven by systematic attrition post-enrolment. For example, if attriting directly admitted female students are stronger than attriting pooled male students and this difference is big enough, the differences in university performance between these two groups could be due to asymmetric attrition rather than genuine differences in admission cutoffs.

However, it does not appear to be the case. First, the data in Table 14 show that, given the data we observe, there are no significant differences in attrition between pooled males and directly admitted females: they are equally likely to attrit, the attrition is equally likely to involve subject change, and, finally, for the students attriting in years 2 and 3, year 1 Cambridge exam performance across the two groups is the same.

Second, we reestimate our main regressions on the sample of *all* first-year students who entered

Table 14: Post-enrolment attrition

	All students		Attrited students		Attrited in years 2 or 3	
	N	% attrit	N	% subject change	N	Year 1 performance
	(1)	(2)	(3)	(4)	(5)	(6)
All subjects						
Directly admitted female	1913	18	349	86	273	-43.7
Pooled male	812	17	141	89	114	-36.5
MI subjects						
Directly admitted female	509	18	91	95	72	-78.3
Pooled male	367	19	69	91	56	-70.4

Note. *% attrit*: Percent of students who accept Cambridge offer but drop out of our sample either because they have no tractable Cambridge exam data or they change subjects. *% subject change*: Of the students who drop out of our sample, percent that experienced subject change. *Year 1 performance*: standardized performance in year 1 exams for students who drop out of our sample in years 2 and 3, but are present in year 1. *MI subjects*: Mathematically intensive subjects: Economics, Engineering and Mathematics. Physical Sciences are not included due to data limitations (more details are in Section 4.1).

in 2013-2016, regardless of whether they drop out in later years.[32] Furthermore, to this we can add two more years of data, the students who entered in 2017 and 2018 for whom we have their year 1 but not year 2 or 3 performance (and hence we did not use them in the main estimations in the paper). What this means is that we are able to include students who will subsequently drop out of the sample in years 2 and 3 (when, in aggregate, most of attrition occurs) Table 15 reports gender regressions (columns (1)-(3)) and school-type regressions (column (4)), and shows that our results are unchanged compared to the main results in Tables 6 and 11 respectively. This gives us confidence that the findings in our paper are not driven by post-enrolment attrition.

[32]We are grateful to an anonymous referee for making this suggestion.

Table 15: Performance of all year 1 students

	Gender, all (1)	Gender, MI (2)	Gender, non-MI (3)	School (4)
Pooled male	0.15	0.29	0.06	
	(0.03)	(0.06)	(0.06)	
Pooled privately-funded				-0.15
				(0.04)
Observations	3935	1290	1257	3494
Subject FE	✓	✓	✓	✓
Application college FE	✓	✓	✓	✓

Note. Each column reports the results from a different OLS regression. *Pooled male*: a dummy = 1 of pooled male and = 0 if directly admitted female. *Pooled privately-funded*: a dummy = 1 for pooled candidates from privately-funded schools and = 0 for directly admitted candidates from state-funded schools. Dependent variable: the standardized score obtained in first-year exams. Sample: in column (1), the sample is pooled males and directly admitted females in all subjects; in column (2), the sample is pooled males and directly admitted females in mathematically intensive subjects (except Physical Sciences, due to data limitations, see Section 4.1); in column (2), the sample is pooled males and directly admitted females in non-mathematically intensive subjects (except Biological Sciences, due to data limitations, see Section 4.1); in column (4), the sample is pooled privately-funded UK students and directly admitted state-funded UK students, in all subjects. *Estimation*: OLS. Robust standard errors are reported in parentheses.

D Offer probabilities

Here, we compare the probability of an applicant getting an offer from Cambridge across the two genders, conditional on high school qualifications. This complements the (unconditional) average offer probabilities in Table 2, and feeds into the discussion in Section 5.1 where we contrast the conclusions one might draw from naively comparing offer probabilities to those we arrive at using our method.

We estimate the OLS regression of the probability of receiving a Cambridge offer, conditional on gender and high school qualifications.[33] The results are in Table 16, separately for mathematically intensive subjects (column (1)) and the rest (column (2)).

For mathematically intensive subjects, high school qualifications include general exams, advanced math score and math modules (described in Section 5.6.2 and Appendix H). For the rest of the subjects we only use general exams, since advanced maths is not essential and many applicants

[33]We report OLS for interpretability, but probit regressions give similar results, both qualitatively and quantitatively.

do not have it. Due to data limitations, Physical and Biological Sciences have to be excluded (more details are in Section 4.1).

To summarize Table 16, women are around 5% more likely to be admitted into mathematically intensive subjects (where they are in a minority) and 4% less likely to be admitted into other subjects (where they are in the majority). These differences run in the same direction, and are slightly larger in magnitude than those based on unconditional means in Table 2.

Interestingly, although the offer probabilities alone cannot provide conclusions about trade-offs, the observation that, for both groups of subjects, they are higher for the underrepresented gender is suggestive of the idea that the decisionmaker is interested in gender balance, a key premise of our theoretical model.

Table 16: Offer probability, by subject group

	MI subjects (1)	Non-MI subjects (2)
Male	-0.05	0.04
	(0.01)	(0.01)
General exams	0.07	0.16
	(0.01)	(0.01)
Advanced maths score	0.87	
	(0.03)	
Math modules	-0.34	
	(0.01)	
Observations	6109	4747

Note. Each column reports the results from a different OLS regression. *Male*: a dummy = 1 if male and = 0 if female applicant. Dependent variable: probability of being given offer at Cambridge. Sample: in column (1), applicants to mathematically intensive subjects (Economics, Engineering and Mathematics); in column (2), applicants to the rest of subjects (Law and Medicine). *Estimation*: OLS. Robust standard errors are reported in parentheses.

E Quantifying the size of trade-off

Here we show the derivation underlying the results in Section 5.3 where we quantify aggregate implications of our trade-off estimates in Mathematically Intensive subjects.

We perform two back-of-the-envelope calculations. The first asks if, using the *realized* post-entry performance data, we used the same cutoff for direct admissions from both groups, how much would average performance increase relative to the case where the cutoffs differ by the magnitude of the *lower bound* on $g_1 - h_1$ calculated above. The second calculation asks how many more h-types (and fewer g-types) are admitted because of the difference in cutoffs.

E.1 Performance forgone

For the first exercise, suppose we move from a situation of different thresholds g_1, h_1 to a common threshold c satisfying $h_1 < c < g_1$. Since the space constraint has to be met, c must satisfy

$$N_g F_g(c) + N_h F_h(c) = N_g F_g(g_1) + N_h F_h(h_1) \tag{18}$$

because total number of admits, and therefore total number of rejections, must remain unchanged. Note that the RHS of equation (18) is simply the total number of rejects, which is directly observable. Furthermore, (18) will have a solution by the intermediate value theorem and it will be unique because the LHS of (18) is strictly increasing in c. The solution can be obtained by

$$\min_c \begin{bmatrix} \sum_{i=1}^{N_g+N_h} 1\{Y_i \le c, G_i = g\} + \sum_{j=1}^{N_g+N_h} 1\{Y_j \le c, G_j = h\} \\ -\{N_g F_g(g_1) + N_h F_h(h_1)\} \end{bmatrix}^2 \tag{19}$$

As long as the c solving (18) is larger than g_3 (the cutoff for rejections), one can solve (18) using the performance data of g-types entering from the pool to calculate the terms $1\{Y_i \le c, G_i = g\}$. The terms $1\{Y_i \le c, G_i = h\}$ are observable for h-types because they are supra-marginal, i.e. $c > h_1$.

Once we get c, we can calculate the ex-post performance gain from moving to a common threshold as

$$\sum_{i=1}^{N_g+N_h} Y_i \times \{1\{c \le Y_i \le g_1, G_i = g\} - 1\{h_1 \le Y_i \le c, G_i = h\}\}$$

This gain in average performance can be calculated although g_1 and h_1 are not individually known. That is because we can calculate

$$\sum_{i=1}^{N_g+N_h} Y_i \times \{1\{Y_i \ge c, G_i = g\} + 1\{Y_i \ge c, G_i = h\}\} - Total\ Current\ Performance \tag{20}$$

E.2 Additional females admitted

As for the second exercise, the number of additional h-types equals $N_h[F_h(c) - F_h(h_1)]$ which equals the reduction in g-types $N_g[F_g(g_1) - F_g(c)]$, which can be calculated exactly once c is known from (18) by

$$N_h \left[F_h \left(c\right) - F_h \left(h_1\right)\right]$$

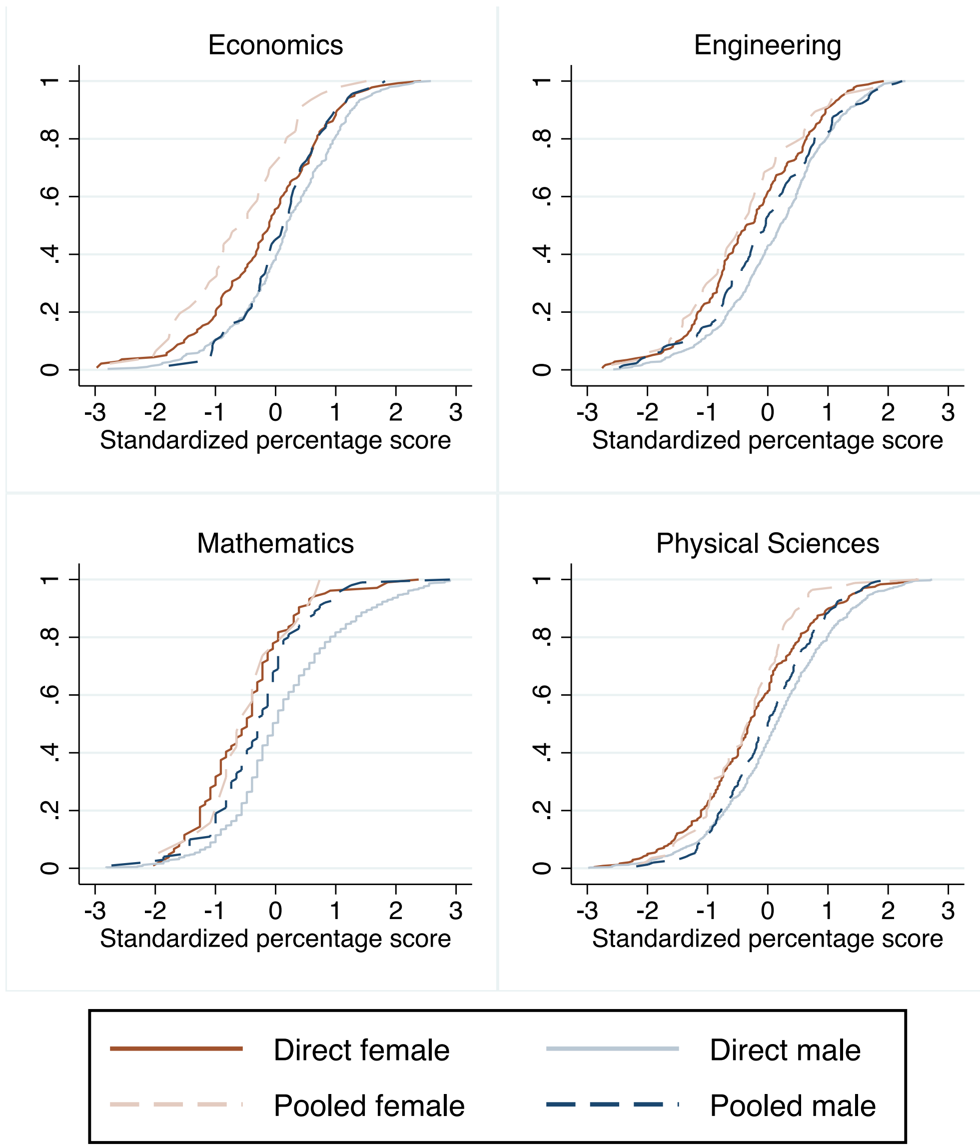
which is the difference between the total observed number of h students admitted directly and how many of them have scored above c in their exams.

F Performance distributions by subject

Figure 3 in Section 5.2 shows that, in mathematically intensive subjects, pooled males have stochastically higher exam scores than directly admitted females *throughout the distribution.* This confirms our finding that the university faces a gender equality performance trade-off in these subjects. Figure 6 below shows that the same also holds in each individual subject in the mathematically intensive group (i.e. Economics, Engineering, Mathematics and Physical Sciences.)

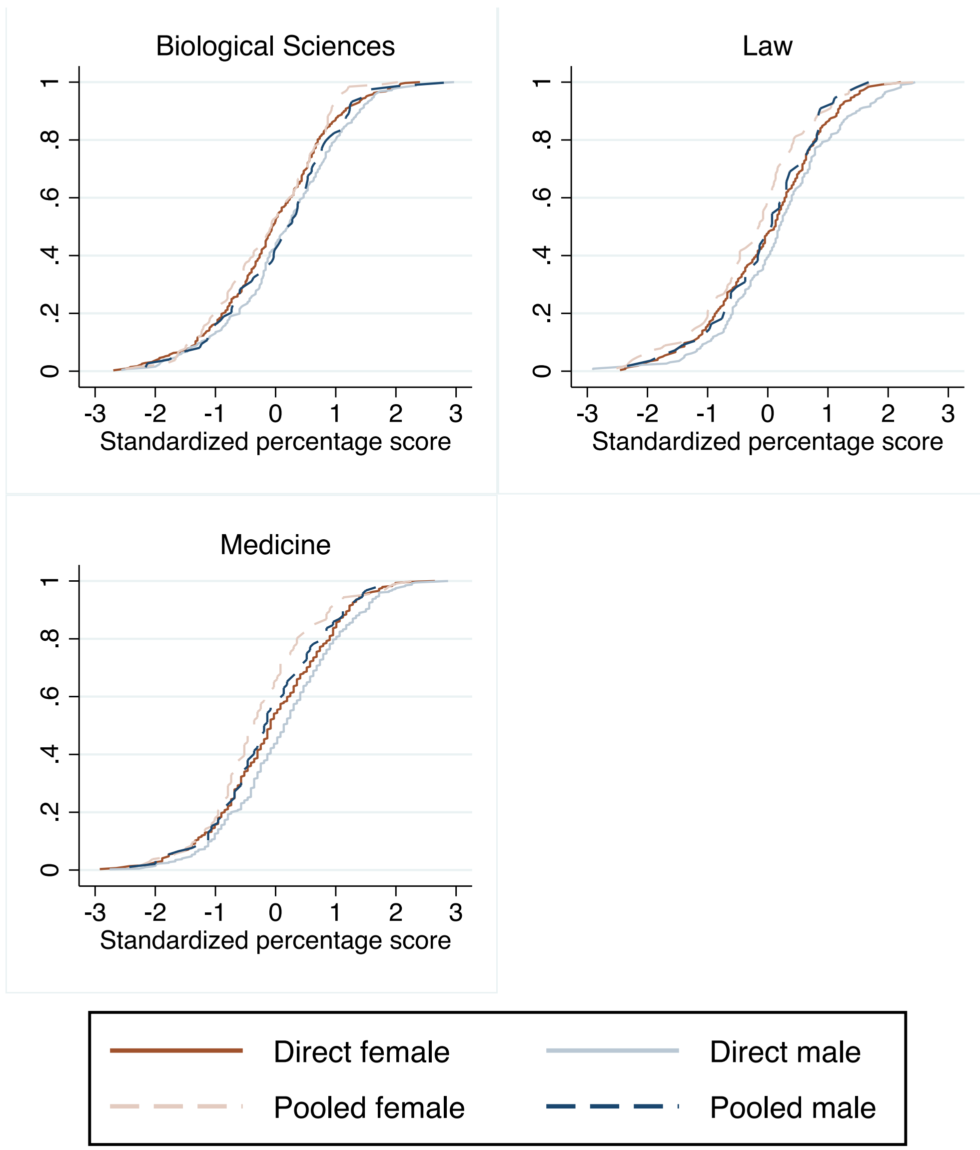
In contrast, in non-mathematically intensive subjects (Figure 3, right graph), the performance of pooled males and directly admitted females is similar, in line with our earlier finding that there is no gender-performance trade-off for these subjects. Below, in Figure 7 we see the same pattern in the individual subjects that comprise this group, viz. Biological Sciences, Law and Medicine.

Figure 6: First-year exam scores by pool status and gender, MI subjects



Note. The graph shows the cumulative distribution function of standardized first-year exam scores for different subgroups of students in the mathematically intensive subjects, by subject.

Figure 7: First-year exam scores by pool status and gender, non-MI subjects



Note. The graph shows the cumulative distribution function of standardized first-year exam scores for different subgroups of students in the non-mathematically intensive subjects, by subject.

G College participation in the pool

In Section 5.4 we discuss the correlation between a college's performance and their participation in the pool. Figure 8 shows that better performing colleges are more likely to put (male) students in the pool.[34] To deal with this threat to identification, we control for application college fixed effects in our regression.

Figure 8: College performance and contributions to the pool

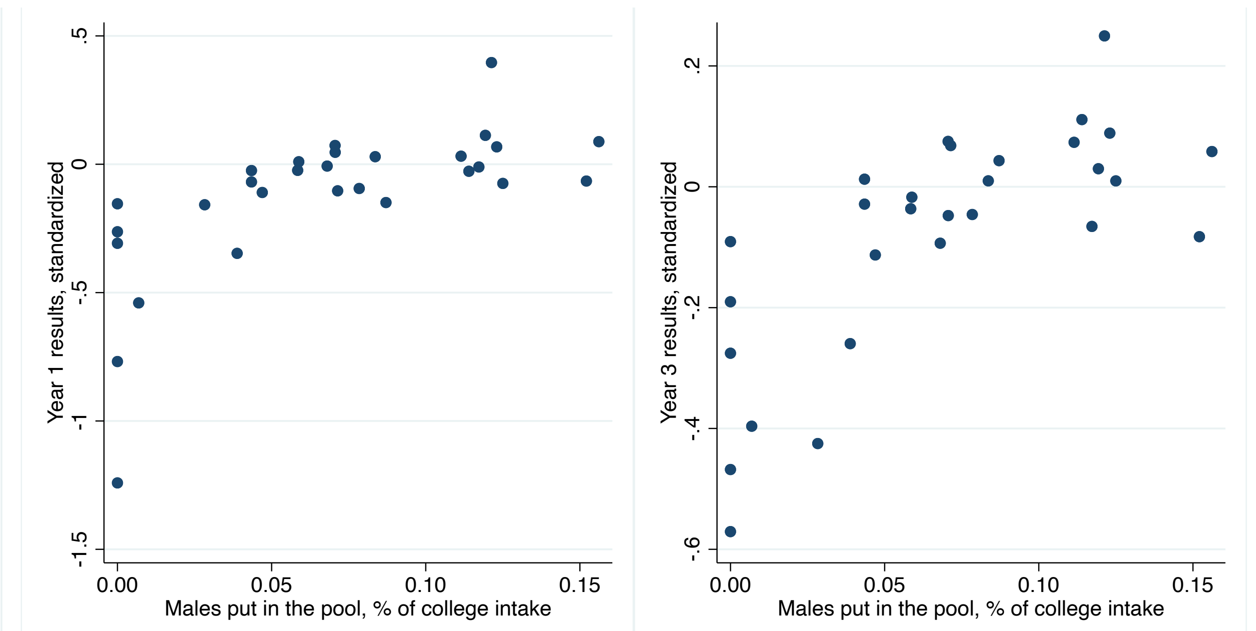
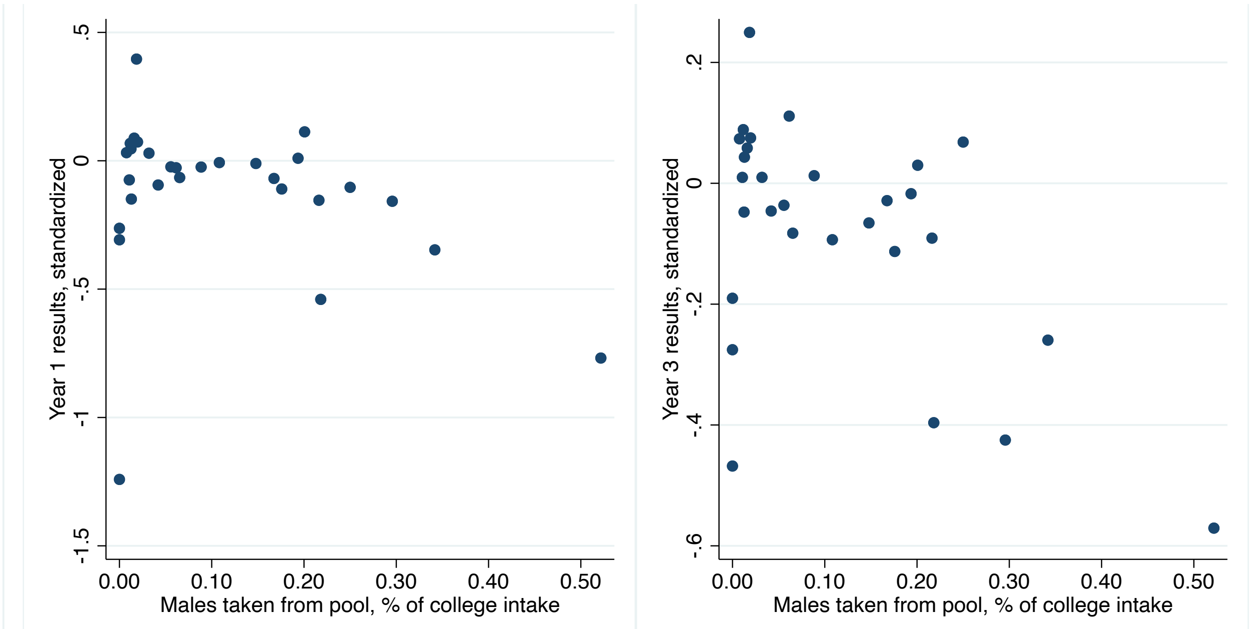


Figure 9 shows that colleges that take more males from the pool, if anything, tend to perform worse. This correlation goes in the opposite direction of the effect that we find, and so this is not a threat to identification. Nevertheless, we show that qualitatively our results are unchanged when controlling for offer college fixed effects which would address this threat had it been present (Table 6).

[34]The patterns in Figures 8 and 9 and hence conclusions are the same if we look at all applicants put in/taken from the pool, rather than just male applicants.

Figure 9: College performance and withdrawals from the pool



H High school qualifications: background and summary statistics

In this appendix, we give the details of high school qualifications obtained under the British exam system, and explain the three measures we focus on in section 5.6.2.

Typically, pupils apply to Cambridge when they are 17 years old, and the application lists their results from the two rounds of British system exams, first taken at 16 (general exams or GCSEs) and second at 17 (advanced exams or AS). Cambridge admission decisionmakers look at both sets of exams, with the advanced exams considered more relevant because they are more recent, advanced and specialist.

General exams: General exams are known as GCSEs in the UK and IGCSE internationally. They are nationally graded, compulsory board exams for all school pupils, whether or not they go on to higher education. They are taken when students are, typically, 16 years old, roughly 16 months before a student applies to Cambridge. Usually, these exams are taken in 10 different subjects and for each, a pupil receives grades A*, A, B,...,G, where A* is the highest and G is the lowest.[35] The scale is quite coarse, and many students applying to Cambridge will have top marks in these exams. For its admission decisions, Cambridge tends to focus on a crude summary statistics from these exams, namely the number of A*s obtained by the candidate in all the subjects. Hence, this is also the measure we use in our analysis to capture general exam performance (and to which we refer as the 'general exams' variable).

[35] There is also 'U', for ungraded.

Advanced exams[36] : In the last two years of their school career in the British system, pupils specialize and study, typically, only three or four subjects. For those wanting to continue to university, the choice of these subjects is informed by the admission requirements of their desired degree. The exams in these subjects give rise to the second set of school qualifications, the advanced exams, known as A-levels. These exams come in two parts because they are spread over two years. The first and earlier part, known as Advanced Subsidiary (AS) levels is the one relevant to our analysis because the constituent exams take place before the students apply to university.[37]

Advanced exams are typically taken in all of the A-level subjects the applicant is studying, four-five months before they apply to Cambridge. Hence, AS levels are the most recent set of school results available to university decisionmakers at the time of the candidates' application to Cambridge, and tend to carry significant weight in the admission process. In contrast with general exams, for advanced exams admission decisionmakers focus on grades obtained in individual subjects, particularly those that are relevant to the university course the candidate is applying for. Since mathematics is the one subject that is key to all of the degrees in our mathematically intensive group, we use the score attained in the advanced AS maths exam ('math score') as a key advanced qualification.

Another important characteristics of the A-level system is that school pupils have some flexibility in the number of AS modules or courses (and hence exams) they choose to enrol in for each subject they study at A-levels. The candidates taking more modules are covering more material and coping with a higher workload. For this reason, we also introduce another variable, which is the number of math modules ('math modules') the candidate has taken. Although we cannot directly observe this, our data contain the maximum points that a candidate can earn in maths across all modules they take. Using the fact that a typical module has 100 points available, we create a proxy for the number of math modules taken by each candidate.

The summary statistics for high school qualifications are in Table 17, which displays the averages of high school qualifications by gender, admission tier and subject type. First, the table shows that, within both admission tiers, males have lower scores than females in general exams, but higher math scores, and they also take more math modules. Second, the gender gap in math preparation is larger for applicants to mathematically intensive subjects. Third, the gender differences persist also when we compare directly admitted females and pooled males.

[36] Advanced exam system underwent a substantial reform in 2017. The data used in this paper are pre-reform, and so the description here refers only to the pre-reform arrangements.

[37] The second part, known as A2, is taken in the summer before starting university, by which time the universities have already made most of their offers. For this reason, Cambridge offers are usually conditional on the final A-level grades (see Section 2.3 and Appendix C.1 for more details).

Table 17: Means of high school qualifications

	MI subjects				Non-MI subjects			
	Direct		From pool		Direct		From pool	
	M	F	M	F	M	F	M	F
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
General exams	-0.14	0.14	-0.24	0.11	0.06	0.20	-0.01	0.17
Advanced Math exams								
Math score	0.44	0.13	0.32	0.05	-0.47	-0.62	-0.45	-0.61
Math modules	6.54	5.88	6.28	5.75	4.18	3.75	4.25	3.82

Note. Sample: Students from schools using British system exams for whom both Advanced and General exam information is available. M is male, F is female. General exams and math score are standardized.
***, **, and * in F columns denote that F is different from M in the same admission tier and subject group at 1%, 5% and 10% level respectively.
***, **, and * in From Pool M columns denote that M is different from Directly Admitted F in the same subject group at 1%, 5% and 10% level respectively. *MI subjects:* Mathematically intensive subjects, viz. Economics, Engineering, Mathematics and Physical Sciences. *Non-MI subjects*: Non-mathematically intensive subjects, viz. Biological Sciences, Medicine and Law.

I High school qualifications: Additional analysis

I.1 High school qualifications: Average gender gap, MI subjects

We ask whether high school qualifications are related to subsequent exam performance in Cambridge, and whether the observed gender differences in these qualifications explain the gender gap in *average* university performance in mathematically intensive subjects.

To do this, we focus on students directly admitted to mathematically intensive subjects, and regress their first-year exam performance on gender dummy alone (Table 18, column (1)), then include high school qualifications (column (2)) and finally interact high school qualifications with gender (column (3)). The results should be interpreted with caution, as they are conditional on being admitted to the university.

As discussed in Section 5.6.2, all three high school qualifications are significantly positively correlated with performance at Cambridge. Amongst them, math score has three times the coefficient of general exams (both standardized). Whilst simply adding these qualifications does not affect the size of the average gender gap, when we interact them with gender, some of the interaction terms are significant and the gender gap shrinks (column (3)). This suggests that (a) males and females have a different relationship between high school qualifications and performance at university and (b) some of the *average* gender gap observed is explained by these qualifications.

This is in contrast with our finding in Table 9 where we saw that *on the margin* we are investigating, i.e. pooled males vs directly admitted females, controlling for the qualifications (and allowing for gender heterogeneity in their effects) does not reduce performance gap between the two.

Table 18: Performance, gender and high school qualifications of directly admitted candidates, mathematically intensive subjects

	(1)	(2)	(3)
Male=1	0.40	0.39	0.22
	(0.05)	(0.05)	(0.16)
General exams		0.09	0.14
		(0.02)	(0.05)
Math score		0.32	0.40
		(0.03)	(0.05)
Math modules		0.05	0.03
		(0.01)	(0.02)
Male=1 × General exams			-0.06
			(0.05)
Male=1 × Math score			-0.10
			(0.06)
Male=1 × Math modules			0.03
			(0.02)
Observations	2048	2038	2038
Subject FE	✓	✓	✓
Application college FE	✓	✓	✓

Note. Sample: All directly admitted applicants who sat British system school exams and for whom both general and advanced exam information is available.

I.2 Oxaca decomposition: Marginal gender gap, MI subjects

We now ask whether differences in *marginal* performances, i.e. between pooled males and directly admitted females in mathematically intensive subjects can be explained by high school qualifications. To complement our analysis in Table 9, we perform Oxaca decomposition and report the results in Table 19.[38]

[38] We are grateful to an anonymous referee for this suggestion.

The decomposition is based on regressing performance of (i) pooled males and (ii) directly admitted females separately on high school qualifications.[39] That is, we decompose the difference in mean between pooled males and directly admitted females as

$$
\begin{aligned}
E\left(Y|G=g\right)-E\left(Y|G=h\right) &= \underbrace{\left(\gamma-\eta\right)' E\left(X|G=g\right)}_{\text{Productivity-effect at mean } X \text{ of group } g} + \eta' \underbrace{\left[E\left(X|G=g\right)-E\left(X|G=h\right)\right]}_{\text{Covariate-effect}} \quad (21) \\
&= \underbrace{\left(\gamma-\eta\right)' E\left(X|G=h\right)}_{\text{Productivity-effect at mean } X \text{ of group } h} + \gamma' \underbrace{\left[E\left(X|G=g\right)-E\left(X|G=h\right)\right]}_{\text{Covariate-effect}} \quad (22)
\end{aligned}
$$

where for $i = \{g, h\}$, such that g is pooled males and h is directly admitted females, Y_i is standardized first-year Cambridge exam performance, X_i is the vector of the high school qualifications (general exams, math score and math modules), γ is the vector of coefficients on X obtained in regression of Y on X for subgroup g, and η is the vector of coefficients on X obtained in regression of Y on X for subgroup h.

The results in Table 19 confirm our earlier findings in Table 9. Although there are differences in the relationship between covariates and performance across the two groups (the 'productivity-effect'), they run in the opposite direction of the observed university performance gap, and so does the much smaller effect resulting from the difference in levels of covariates. Hence, on the margin that we are investigating, the gap between directly admitted females and pooled males in mathematically intensive subjects does not appear to be explained by high school qualifications. This is despite the fact that these qualifications are correlated with performance and partly account for the *average* gender gap in mathematically intensive subjects (see Appendix I).

[39] These equations are the same as those in Table 9, except they (i) estimate the effects of regressors separately using two regressions rather than via interaction terms in one combined regression and (ii) exclude fixed effects.

Table 19: Oxaca decomposition: Gender in MI subjects

	Eq-n (18)	Eq-n (19)
	(1)	(2)
Lower bound $\mathbb{E}[X_g] - \mathbb{E}[X_h]$	0.163	0.163
Difference in intercepts (g-h)	0.442	0.442
Productivity effect	-0.251	-0.263
Covariate effect	-0.016	-0.029

Note. *Sample*: pooled males and directly admitted females, enrolled in mathematically intensive subjects, who sat British system school exams and for whom both general and advanced exam information is available. Column (1) reports Oxaca decomposition based on equation (21) above and column (2) reports Oxaca decomposition based on equation (22) above. g are pooled males, and h are directly admitted females. Covariates (pre-admission characteristics) are general exams, math score and math modules. *Difference in intercepts*: The difference between the intercepts of the regression equation of Y_g on X_g and of that of Y_h on X_h (i.e. the gender gap unexplained by the covariates.) Note that the first two rows of the table are the same because the intercepts and the lower bound of differences in covariates do not differ between equations (21) and (22).

I.3 Relative weights on qualifications, MI subjects

We now look into whether when making direct admission decisions, admission officers may be underweighting maths relative to other qualifications.[40] To do this, we ask whether the weights they assign to the individual qualifications are different from the weights with which these qualifications subsequently enter Cambridge performance equation. Although there is the selection problem, the two tier admission system allows us to partly circumvent it.

Specifically, we estimate two equations: first, we regress first-year exam performance at Cambridge on high school qualifications of all applicants admitted in mathematically intensive subjects. The results of this are reported in Table 20, column (1).

Second, for the same group, we regress the probability of being admitted directly, i.e. in the first tier, on the same high school qualifications. The results of this are reported in Table 20, column (2).

For each regression, we compute the ratio of (i) the coefficient on math score over the coefficient on general exams, and (ii) the coefficient on math module over the coefficient on general exams. For

[40] We are grateful to an anonymous referee for this suggestion.

(i) they are 2.1 and 5.5, and for (ii) they are 0.7 and 1, respectively, i.e., for both math variables, the weight relative to general exams is higher in the performance equation than in the admission equation (roughly 200% and 40% higher, respectively). This suggests that admission officers are not putting sufficient weight on maths qualifications, relative to general exams, when making admission decisions.

Table 20: Role of high school qualifications, mathematically intensive subjects

	Probability of direct admission (1)	Year 1 performance (2)
General exams	0.01	0.06
	(0.01)	(0.02)
Math score	0.03	0.32
	(0.01)	(0.02)
Math modules	0.01	0.06
	(0.00)	(0.01)
Observations	2504	2504
Subject FE	✓	✓
Application college FE	✓	✓

Note. Sample: Students enrolled in mathematically intensive subjects, who sat British system school exams and for whom both general and advanced exam information is available. *Dependent variables:* Column (1) a dummy = 1 if directly admitted, = 0 if admitted through the pool; Column (2) first-year exam performance, standardized. Estimation: OLS.

I.4 Counterfactual performance and admission cutoffs

Using estimates in Table 9 we construct the counterfactual predictions of first-year exam performance in Cambridge for all applicants to mathematically intensive subjects.[41] Figure 10 plots the predicted performance CDFs for all female applicants and all male applicants that ended up in the pool. [42] Horizontal lines show fractions that were admitted from the two groups. We see the same pattern with the predicted results as with our regressions using actual performance: the implied performance cutoff (-0.6) is lower for directly admitted females than for pooled males (-0.2).

[41]We are grateful to an anonymous referee for suggesting this.

[42]This excludes Physical Sciences due to data limitations (more details are in Section 4.1.

Figure 10: Distributions of predicted performance