unibz
Fakultät für Wirtschaftswissenschaften
Facoltà di Economia
Faculty of Economics and Management

BEMPS –
Bozen Economics & Management
Paper Series

NO 117/ 2026

Patterns in University Applications: Socioeconomic Status, Gender, and Subject vs. Institution Preferences

F. Hertweck, R. Maris, M. Tonin, M. Vlassopoulos

Patterns in University Applications: Socioeconomic Status, Gender, and Subject vs. Institution Preferences*

Friederike Hertweck†, Robbie Maris‡, Mirco Tonin§, Michael Vlassopoulos¶

January 11, 2026

Abstract

This paper examines university application patterns in the UK, focusing on the joint decision of selecting both an institution and a subject. Using administrative data from the Universities and Colleges Admissions Service (UCAS) covering almost all undergraduate applications between 2008 and 2021, we document three key facts: (i) students generally choose subject before university: they apply on average to around 1.6 subject areas across 4.6 institutions, and roughly half apply to a single field across multiple universities; (ii) there are significant gender gaps in application and offer rates that reflect field composition; (iii) high-socioeconomic status students submit more applications, apply less to local institutions, and obtain more offers, but these differences shrink sharply once we control for attainment and the selectivity of the programmes that students apply to. An expert survey suggests that several of these patterns run against conventional wisdom.

Keywords: Higher Education; Application patterns; UCAS data; Gender; Socioeconomic Status

JEL Codes: I20, I23, M38

*We would like to thank workshop and conference participants at Humboldt University, UCL, Linz, Trento (22nd Brucchi-Luchino Labor Economics Workshop), and the workshop on the Economics of Education: EdTech and Gender. The survey received ethical approval from the University of Southampton [Ethics/ERGO no: 93571]. Rachel Kühn and Valeria Ride provided excellent research assistance. Friederike Hertweck acknowledges the receipt of travel funds provided by the Fritz Thyssen Foundation. This work was undertaken in the Office for National Statistics Secure Research Service using data from ONS and other owners and does not imply the endorsement of the ONS or other data owners. Refine.ink was used to check the paper for consistency and clarity. Corresponding author: Friederike Hertweck - friederike.hertweck@rwi-essen.de.

†RWI – Leibniz Institute for Economic Research

‡University College London

§Free University of Bozen-Bolzano, FBK-IRVAPP, IZA, CESifo and Dondena Centre

¶University of Southampton, IFS, and IZA

1 Introduction

University attendance shapes both individual outcomes, such as earnings and career trajectories (Altonji, Arcidiacono, & Maurel, 2016; Patnaik, Wiswall, & Zafar, 2021; Lovenheim & Smith, 2023) and broader societal outcomes, such as economic inequality and mobility and economic growth (Goldin & Katz, 2009).[1] A better understanding of how students make university choices, both what to study (subject) and where to study (institution), is important for designing policies that can improve efficiency and equity in access to higher education. This study examines how students navigate this joint decision process in the United Kingdom, where applications to higher education involve a simultaneous selection of both an institution and a subject area through the centralized Universities and Colleges Admissions Service (UCAS) system.

University choice involves two distinct but interconnected decisions: what academic institution to attend and what program/subject area to study. In the UK and in many other countries, this decision is joint, as students select a combination of University and program, through centralized application systems like UCAS or in a decentralized manner. In other countries, such as the US, the decision is instead sequential: first students choose University and then major, whilst at University (Bordon & Fu, 2015). Nonetheless, it is possible that even in such systems, students may prioritize the subject they want to study—such as Engineering or Medicine—and then choose a university based on its strengths in that field. Thus, even in seemingly sequential choice systems, students may still prioritize their field of study when making decisions. In the context like the UK where the choice is joint, it is unclear whether students typically decide to study a particular field (e.g., Medicine) and then identify suitable universities, or they begin with a preferred institution (e.g., Oxford) and choose among its available programs.

We ask a simple question that cuts across systems: do applicants build portfolios subject-first or institution-first, and how does this vary by gender and socioeconomic status (SES)? This distinction is important because institution and subject choices are influenced by different factors. The decision to apply to a University may be influenced by the University's status or location, whereas the choice of subject may be driven by personal interest, aptitude or career expectations (Wiswall & Zafar, 2015; Delavande & Zafar, 2019; Broecke, 2015). Clarifying this decision-making order is essential for designing admissions policies that aim to widen participation and improve match quality.

Another important dimension that we address in this paper is the quality of choices students make: whether they apply to more or less selective institutions and programs relative to their academic record and whether this varies by gender and socioeconomic status. Understanding how application patterns relate to the prestige or value of programs, as captured by selectivity, is critical for evaluating efficiency and equity in the system.

This paper provides insight into these issues drawing on millions of applications to UK universities from UCAS (UCAS, 2022). Our data cover almost the universe of applications to UK higher education institutions between 2008 and 2021. These data allow us to document detailed patterns that are difficult to uncover with survey data alone. Moreover, we provide evidence throughout the pre-enrollment process, spanning both applications and admissions.

Beyond better understanding how students structure their choices, examining application behaviour is crucial for assessing access to tertiary education. Differences in application patterns by gender and socioeconomic status can reflect broader disparities in opportunities and barriers to higher education. By documenting these differences, our study contributes to the discussion on equity in university admissions and informs policies aimed at expanding access.

We document three key facts: (i) within the 5 choices allowed by UCAS, students on average choose 1.6 fields across 4.6 universities, with roughly half of applicants applying to a single field across multiple

[1] See Britton et al. (2022), for recent evidence on the returns to different university degrees in the UK.

institutions, thus indicating that they mainly choose subject area and then university when constructing their application portfolios; (ii) there are some gender differences in the application process, in particular, males send more applications and, conditional on sending the same number of applications, receive about 7% more offers. These gaps do not vanish after conditioning on demographics, A-level attainment, and portfolio tier but are largely explained by differences in subject composition rather than institution selectivity or discriminatory treatment; (iii) there is a significant socioeconomic gradient in the application process, with high socio-economic status (SES) students submitting more applications. Even when we condition on submitting the maximum number of applications, i.e., 5, high SES students apply to more institutions and fewer in their home region, and receive more offers than low SES students. This pattern suggests that students with higher-SES may have advantages in navigating the application process. However, these differences shrink markedly once we control for attainment and the selectivity of portfolios, leaving small residual differences.

In addition to documenting patterns in application behaviour, we compare our findings to the views of experts in the field of higher education, elicited through a survey. This expert survey offers an external benchmark, enabling us to assess whether the observed patterns align with what experts expect. As a second benchmark, we use ChatGPT, a widely used large language model trained on publicly available information, to generate predictions about application patterns. These comparisons offer insights into the degree of alignment (or misalignment) between observed student behaviour and both expert understanding and model-based inference of the university application process. Experts' views are consistent with the observed patterns on SES gaps, though they underestimate the number of universities students apply to, and they expect females to submit more applications and receive more offers.

These findings have important implications for admissions policies. If students prioritize subject over institution, efforts to expand access should account for how subject availability influences application behaviour. For instance, if high-demand subjects are concentrated in a few institutions, students with financial or personal constraints on mobility may face limited choices. Furthermore, the gender and socioeconomic disparities we document highlight potential barriers that may require targeted interventions to ensure equitable access to higher education. In addition, understanding that applicants may first choose a field of study and then an institution can inform the design of matching mechanisms that better align with actual decision-making processes. Indeed, Calsamiglia, Fu, and Güell (2020) emphasize the importance of accounting for the structure of applicants' preferences in school choice markets. When preferences are hierarchical (for example, subject-first), mechanisms that reflect this hierarchy in how options are ranked can reduce incentives for strategic misranking relative to a single flat list. Therefore, designing university admission mechanisms that recognize such hierarchical preferences may lead to more efficient and equitable outcomes by limiting the scope for strategic manipulation.

Our study contributes to two strands of the literature. First, it extends research on university choice. Prior work has examined the role of expected returns to different majors, showing that students respond to both expected and actual earnings differentials across fields (Beffy, Fougere, & Maurel, 2012; Arcidiacono, Hotz, & Kang, 2012; Stinebrickner & Stinebrickner, 2014; Wiswall & Zafar, 2015). Other studies have focused on the roles of non-pecuniary considerations and financial constraints (Delavande & Zafar, 2019). Institutional characteristics also play a role, with evidence from the UK showing that student satisfaction, university rankings, and subject-specific league tables influence application behaviour (Gibbons, Neumayer, & Perkins, 2015; Broecke, 2015; Chevalier & Jia, 2016). In addition, recent evidence points to social context and family also playing an influential role (Altmejd et al., 2021; Barrios-Fernández, 2022), as well as economic conditions (Blom, Cadena, & Keys, 2021). Our analysis complements these studies by providing large-scale evidence from administrative data on how students structure their choices across subjects and institutions.

Second, our study relates to research on how socio-economic background, gender, and other demographic factors influence access to higher education. In the UK, previous studies have shown disparities in access by socioeconomic status (Gibbons & Vignoles, 2012; Belfield, Boneva, Rauh, & Shaw, 2020), while such patterns have also been shown in the US (Hoxby & Avery, 2012). Gender differences in subject choices have also been documented, with women more likely to apply to fields with lower expected earnings despite similar academic ability (Zafar, 2013; Delaney & Devereux, 2019; Sloane, Hurst, & Black, 2021; Campbell, Macmillan, Murphy, & Wyness, 2022). We contribute to this literature by documenting disparities in application patterns and offer rates, as well as by comparing these patterns to expert expectations.

Furthermore, a growing body of research has explored the sequence in which students make decisions regarding what and where to study, with varying assumptions about the order of these decisions. For instance, Pigini and Staffolani (2015), using data from Italy, impose that students first select their field of study and subsequently choose a particular institution. This sequential structure aligns with our findings, as students in our setting exhibit a preference for selecting a field of study. In contrast, other studies suggest alternative sequences that are partly shaped by the specific institutional settings in which they are situated. For instance, Declercq and Verboven (2015) highlight that language barriers in Belgium require students to select the region first and field and/or institution second. Schmidt, Ortúzar, and Paredes (2019) model the university choice in Chile. In one of their models they assume that students consider similarly ranked universities as reasonable alternatives, irrespective of field of study. Yet, they also show that distinct groups of students exhibit different decision-making processes which cannot be clearly predefined. They conclude that the university choice is largely influenced by students' high school performance rather than their socioeconomic status. Our results contribute to this literature by providing direct evidence on the relative importance of field versus institution in a context where the choice is formally joint.

The paper proceeds as follows. Section 2 describes the institutional setting and the data. Section 3 outlines a simple conceptual framework that guides the analysis. Sections 4 and 5 present the main results on application and admission patterns, and Section 6 provides complementary evidence from expert and AI expectations. Section 7 concludes.

2 Institutional setting and data

In this section, we first outline the application and admission process of UK universities. We then provide details on the administrative dataset of applications and admissions. Finally, we describe the survey that we have designed and administered to gauge experts' predictions.

2.1 Application and admission process

The UK has a centralized admission system, managed by the Universities and Colleges Admissions Service (UCAS).[2] The vast majority of UK universities and colleges use UCAS services to manage the application process, therefore, almost all students planning to study for an undergraduate degree in the UK must apply through UCAS. In order to apply, students are required to use UCAS online platform to submit a single application through the so-called main scheme. This process entails registering for the

[2]UCAS is a privately held company that provides educational support services in the UK. The company's main role is to operate the application process for UK universities and colleges and has been formed by the merger of the former admission systems for UK universities and polytechnics.

service, providing personal information, crafting a personal statement[3], and selecting up to five courses for application, with no specific order of preference. Subsequently, applicants must remit an application fee and secure a reference before submitting their application online by the designated deadline. Until 2023, the application fee amounted to 22.50 GBP for a single application (i.e., one course) and 27.50 GBP for multiple applications (i.e., two to five courses).[4]

After receiving the applications, UCAS forwards them to the universities and colleges selected by the applicants. Upon receipt, these institutions adhere to their own internal procedures, which may involve conducting interviews, before determining whether to send an admission offer back to the applicants. In sending offers, an institution can either send an *unconditional offer*, guaranteeing the applicant a place, or a *conditional offer*, contingent upon the applicant meeting specified grades.

An applicant can thus receive up to five offers. They then select up to two options via UCAS: one as their *firm choice* and one as their *insurance choice*. The firm choice indicates that, upon meeting the required grades, the student's conditional offer will be affirmed. The insurance choice acts as a backup in that if the applicant fails to meet the conditions of the firm choice, they will gain admission to the insurance choice university, provided they meet the specified conditions, if any. Of course, an applicant can only designate an insurance choice if their firm choice is a conditional offer. The remaining offers are automatically declined. After receiving the final examination results, students learn whether they have met their "firm" and "insurance" offer conditions.

As illustrated in the timeline of the application process depicted in Figure 1, if some applicants do not receive any offers or decline all their offers, they can apply for additional courses using the so-called Extra. Extra allows students to keep applying, one course at a time, until an offer is secured. Conversely, if an applicant exceeds the condition of their firm offer, they could search for another course via the Adjustment service while retaining their original offer. Since 2022, this adjustment service is no longer available. If students are still unplaced after Extra, or if they missed earlier application deadlines, they can use UCAS Clearing, which enables unplaced students to apply for courses with vacancies directly to the university from July onwards. Thus, Clearing provides a second opportunity to enter higher education, but the number of available places is not known in advance. Finally, students can bypass UCAS entirely by using the Record of Prior Acceptance (RPA) service to directly apply to a course at a university. As we show in the next section, despite the existence of these multiple routes, UCAS main scheme is by far the dominant one in terms of application volume.

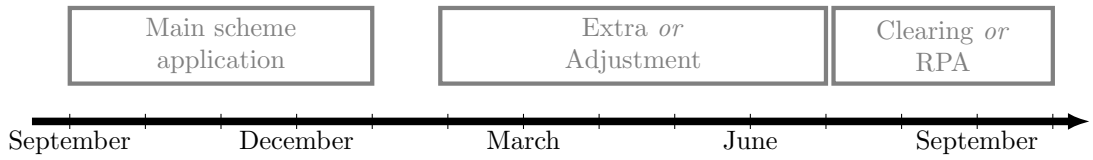


Figure 1: Timeline of application process

2.2 Data and sample

For this study, we use UCAS application and admission data (UCAS, 2022). The dataset contains individual records for almost the entire universe of applicants to undergraduate courses in the UK from 2008

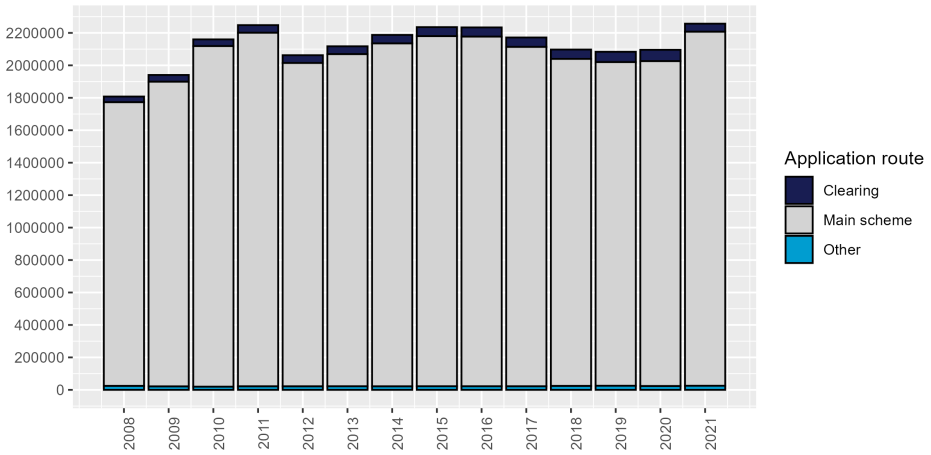
[3]During the period of observation, UCAS personal statements were not field-specific, and applicants had more flexibility in how they wrote them. Recently, UCAS introduced reforms requiring applicants to explicitly mention their chosen subject, with subject-specific guides and clearer instructions for applicants applying to multiple courses.

[4]Since 2024, the single application fee is no longer available and applicants have to pay a fixed fee irrespective of sending just one or up to five applications.

to 2021. These records include demographic, socio-economic, and educational characteristics, neighbourhood information, as well as the application details of 10,550,252 applicants. Overall, we have access to 42,055,924 applications.

The dataset includes both home and overseas students, as well as applications that were incomplete and automatically withdrawn (less than 0.01%). Because we are interested in understanding application behaviour of typical undergraduate students, we restrict the sample to applicants from the UK aged 17 to 25. Figure 2 illustrates the number of applications via the different routes over time. It shows that the vast majority of applications are submitted via the main scheme route: In the first year in our sample, 2008, UCAS received 1,748,585 applications via the main scheme route while only 59,185 applications (3.27%) were submitted via other routes. The figure also shows that there has been a steady increase in the number of applications from 2008 to 2011. From 2012 onward, the tuition fee cap was increased from 3,000 GBP to 9,000 GBP per academic year, leading to a drop in applications in 2012, with a gradual recovery in subsequent years. A more gradual drop is also evident starting in 2017. Throughout the period, the dominant role of the main scheme is clear and, given its preponderance, for the remaining analysis, we concentrate on applications submitted via the main scheme.

Figure 2: Application overall numbers



Notes: The figure illustrates the number of applications sent via the different routes, i.e., Main scheme (96.49% of applications), Clearing (2.44%), RPA (0.61%), Extra (0.42%), and Adjustment (0.03%); the latter three summarised as other.

For each applicant, we observe up to five applications, each corresponding to a course at a higher education institution. The UCAS data provide rich information on applicants and the application process. This includes age, ethnicity, gender, socioeconomic status, high school qualifications (A and AS levels), and region of residence.[5] We also have detailed information about the university application process, observing all applications (subject, course type, institution, route - e.g. main, extra, clearing), offers, and acceptances. We know to which of the 421 Universities and Colleges an application is directed, as well as the specific course, classified according to the Joint Academic Coding System (JACS) into 19 broad fields or 192 narrow fields. In the Appendix, we present the full lists of broad fields (e.g., Engineering and technology) and narrow fields (e.g., Mechanical engineering; Civil engineering; Biotechnology).

More specifically, for UK-domiciled applicants, we observe socioeconomic status, based on the appli-

[5]In addition, the data also include information about the type of high school attended (e.g., grammar, independent) and the region where it is located, as well as the so called "tariff" point system used to summarize achievement in a numerical and comparable format.

cant's parental occupation,[6] using a simplified version of the National Statistics Socio-economic Classification (NS-SEC). We also observe ethnicity summarised into broad groupings (i.e., 'Asian', 'Black', 'Mixed', 'White', 'Other', 'Unknown'), as applicants are asked to record their ethnic origin on the UCAS application form. Providing information on ethnicity and socioeconomic status is voluntary and is not passed to institutions until after the selection process.

To capture socio-economic background, we can use two further indicators, based on the characteristics of the postcode declared by the applicant at the time of their application. In particular, Polar4 classifies small areas across the UK into five groups according to their level of young participation in higher education.[7] Each of these groups represents around 20 per cent of young people and is ranked from quintile 1 (areas with the lowest young participation rates, considered as the most disadvantaged) to quintile 5 (highest young participation rates, considered most advantaged). Furthermore, the Index of Multiple Deprivation (IMD) identifies small area concentrations of multiple deprivation across all of the UK, providing a relative measure of deprivation among small areas (lower layer super output area).[8]

2.3 Ability Measures

We also develop a series of indicators and variables to capture and account for ability in our analyses. We develop both an overall tariff score - which accounts for all qualifications and associated grades applicants hold - and a series of dummies for the ten most popular A levels and whether students achieved an A or higher in each subject. The tariff point scores are derived from the UCAS tariff points system, which is widely used in admissions processes.

To calculate applicants' tariff points, we used a lookup from the HESA dataset that translates qualification names and grades into tariff points.[9] Almost 75% of individual qualifications were matched. Some did not match because the grade was not associated with any points (i.e., the grade was a fail). For these cases, we manually set the tariff points to zero. Some did not match because the qualification names contained discrepancies between the lookup and the UCAS data. For these instances, where more than 1,000 students had such a qualification (over the 14-year sample period), we manually recoded qualification names. This left around 3.5% of unmatched qualifications in which the grade was blank or the qualification name was missing. We excluded individuals who were missing data for more than 25% of their listed qualifications (this results in us dropping 6% of the sample). We winsorized the tariff point score at the 99th percentile.

To go alongside our tariff points variable, we created a series of A level indicators (the most common entry qualification in the UK). We have developed indicators for whether applicants hold any of the ten most popular A levels, including: English, Maths, Further Maths, Economics, Chemistry, Biology, History, Geography, Psychology and Sociology. We also have associated indicators flagging whether students achieved a grade A or higher in each of these A levels. Finally, we have created a categorical variable denoting the number of A levels an applicant holds (0, 1, 2, 3, 4, 5+).

2.4 Selectivity Measures

An important dimension of the application process is the selectivity of the courses applicants apply to. We calculate a course k's selectivity θ_{kjt} as:

[6] Or the occupation of the person contributing the highest income to the household if the applicant is aged 21 years or over.

[7] https://www.officeforstudents.org.uk/data-and-analysis/young-participation-by-area/about-polar-and-adult-he/

[8] https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019

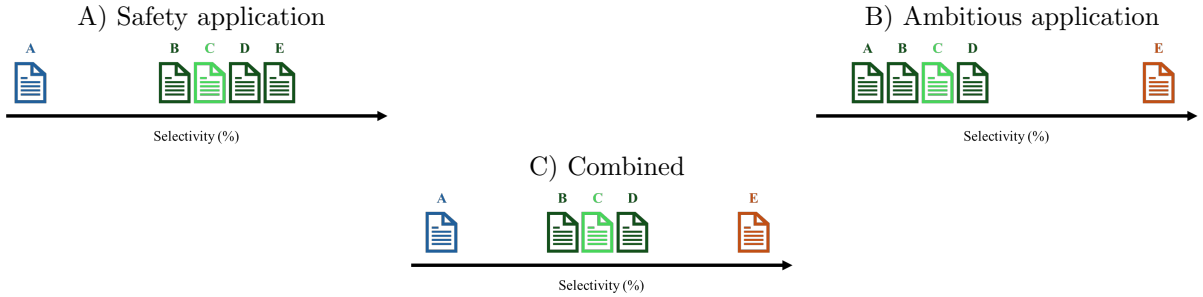
[9] https://www.hesa.ac.uk/collection/c21054/derived/xtpoints

$$\theta_{kjt} = 1 - \frac{\xi_{kjt}}{A_{kjt}} \tag{1}$$

where ξ_{kjt} is the total number of offers for course k at institution j at time t and A_{kjt} is the total number of applications. Higher values of θ_{kjt} indicate more selective courses and $\theta_{kjt} \in [0, 1]$. For instance, a course where $\theta_{kjt} = 1$ is a course where no applicants receive an offer and $\theta_{kjt} = 0$ is a case where all applicants receive an offer.

For each student i, we estimate four selectivity outcome variables: median selectivity, highest selectivity, lowest selectivity and the range of selectivity across individual i's applications. These variables are only calculated for the sample of students who submit five applications. These four moments allow us to explore strategic application behaviour. Having measures of the highest and lowest selectivity for individual i allows us to investigate whether certain groups of applicants are more likely to submit safety applications (applications to courses that are easy to get into) or ambitious applications relative to their median application. The range also tells us whether certain students are more likely to submit a wider spread of applications relative to other groups. This type of strategic behaviour, which we can capture with these variables, is shown in Figure 3. Panel A) depicts a student submitting one safe application, panel B) shows a student submitting one ambitious application and panel C) shows a situation where a student submits both a safe and an ambitious application.

Figure 3: Examples of strategic application behaviour



Notes: In all panels, application C is the application with the median selectivity. In panel A, application A represents a safety application. In panel B, application E represents an ambitious application.

2.5 Expert survey

In line with DellaVigna, Pope, and Vivalt (2019), we conducted a short survey among UK academics and experts in the Economics of Education field to relate our findings to expert expectations about student application patterns.

The five-minute online survey, reported in the Appendix, was conducted in June and July 2024 to two populations: education economists, that is, all the corresponding authors of articles published in the *Economics of Education Review* between 2019 and 2024, plus scholars who participated in the annual Workshops on Applied Economics of Education organized by IWAEE (International Workshop on Applied Economics of Education), and academics based in the UK, that is, 5,000 randomly chosen researchers from 11 UK universities.[10]

In total, we received 124 responses, roughly balanced between the two groups. Table 1 reports descriptive statistics for the two samples, as well as for the full pool of respondents. As expected, the sample of education economists consists almost entirely of economists (97%), mostly based outside of

[10]University College London, University of Birmingham, University of Leeds, University of Ulster, Cardiff University, Newcastle University, Coventry University, University of Hertfordshire, Anglia Ruskin University, and Manchester Metropolitan University.

the UK, and 70% are not directly involved in undergraduate admissions. On the other hand, the UK academics sample consists primarily of academics based in the UK (98%), with more than half involved in undergraduate admissions in the UK, coming from a variety of disciplines. Thus, we have two relevant samples: academic economists who conduct research in education and academics based in the UK.

The first four survey questions asked respondents their opinion about the number of applications, the number of fields and institutions students apply to, and the number of offers applicants receive on average. The next four questions addressed perceived differences between genders and between students from high and low socio-economic background in terms of both number of applications and offers received. Finally, we asked the four demographic questions reported in Table 1. We included in the questions some context, to make sure that every respondent was aware of the relevant institutional features. For instance, before asking about the number of acceptances, we explained the three different responses applicants can receive. Before asking about differences by socio-economic status, we defined what is meant by low and high socio-economic background. Moreover, with the obvious exception of the three questions concerning the number of applications, we asked to consider only those applicants who submitted five applications - the modal case in the dataset - so that differences in the reply, for instance on the number of acceptances, do not derive from different expectations about the number of applications submitted by students. We also instructed respondents to answer with the 2008–2021 period in mind.

Table 1: Descriptive statistics of survey respondents

	Pooled	**Education Economists**	**UK Academics**
Gender:			
Share female	0.47	0.39	0.54
Share male	0.52	0.61	0.44
Discipline:			
Economics	0.48	0.97	0.00
Sociology	0.02	0.02	0.02
Education	0.02	0.00	0.05
Other Social Sciences	0.13	0.00	0.25
STEM	0.26	0.02	0.49
Other	0.10	0.00	0.2
Involved in undergraduate admission:			
Yes, in UK	0.31	0.05	0.57
Yes, elsewhere	0.14	0.25	0.03
No	0.55	0.70	0.40
Affiliation:			
Acad. in the UK	0.58	0.16	0.98
Acad. in Europe (w/o UK)	0.27	0.54	0.02
Acad. in the US	0.09	0.18	0.00
Acad. elsewhere	0.04	0.08	0.00
Other	0.02	0.03	0.00
N	124	61	63

Note: Shares do not add up to 1 if respondents preferred not to answer the specific question.

3 Conceptual framework

In this section, we develop a simple framework to clarify the trade-offs facing UCAS applicants and to guide interpretation of the application patterns we document.

We assume applicants have separable preferences over subjects and institutions:

$$U_i(s,j) \;=\; \alpha_i\, V_i^S(s) \;+\; \beta_i\, V_i^I(j)$$

where $V_i^S(s)$ aggregates subject-specific returns (expected earnings, interest, ability match) and $V_i^I(j)$ aggregates institution-specific returns (quality/prestige, location, costs). The weights $\alpha_i, \beta_i \geq 0$ reflect the relative importance of "what to study" (subject) and "where to study" (institution) and may vary with gender, SES, and ability. A larger α_i/β_i corresponds to *subject-first* prioritization, whereas a smaller ratio to *institution-first*.

What might raise α_i/β_i? A key mechanism operating for many applicants is that they may have more precise information about subjects (A-levels, teacher advice, observed field returns) than about institutions (noisy rankings, limited visits). When facing such information asymmetry, a natural strategy is to commit first to the dimension about which one has better information, while retaining flexibility over the dimension with greater uncertainty.

With at most five slots, applicants hedge between "reach" (more competitive) and "safety" (less competitive) choices.[11] We summarize within-portfolio risk with four moments measured on the five-application sample: (i) the *median selectivity* $\tilde{\theta}_i$; (ii) the *range* $r_i = \max_\ell \theta_{i\ell} - \min_\ell \theta_{i\ell}$; (iii) the *maximum* selectivity $\theta_i^{\max} = \max_\ell \theta_{i\ell}$ (ambition at the top); (iv) the *minimum* selectivity $\theta_i^{\min} = \min_\ell \theta_{i\ell}$ (safety at the bottom).

Observed differences by gender and socioeconomic status (SES) can arise from three broad mechanisms: (i) *Preferences.* Groups may differ in the value placed on subjects and institutions (e.g., field mix, willingness to travel), generating different portfolios even absent constraints; (ii) *Information and beliefs.* Groups may differ in beliefs about admission chances or returns, or in application know-how (e.g., hedging strategies). Information gaps can produce under- or over-reach relative to attainment; (iii) *Constraints* Financial or geographic constraints, school-level counseling differences, or unobserved application quality can affect both portfolios and offers. Our data do not allow one to distinguish between these different mechanisms.

To guide interpretation, we highlight three stylised implications:

- **Subject-first portfolios:** When α/β is high, applicants choose multiple institutions offering their preferred field(s), resulting in portfolios with few distinct subjects and high institution variety.
- **Hedging under uncertainty:** Applicants facing admissions risk spread applications across the selectivity spectrum. Applicants with limited information, lower confidence, or higher perceived costs may apply to fewer courses or cluster in safer options.
- **Group differences in preferences or beliefs:** Differences in application patterns across gender or SES may reflect variation in α/β, in beliefs about admission chances, or in constraints (information, cost, or aspirations). For example, female applicants may pursue high-barrier fields with constrained capacity (e.g. medicine), yielding fewer offers even with similar attainment. Conversely, high-SES applicants may submit broader or riskier portfolios due to stronger guidance or higher application confidence.

[11] UCAS submissions are not ranked at submission, and universities do not observe an applicant's other choices, so applicants can mix reach and safety without incurring preference-signaling costs that exist in systems requiring ranked applications.

4 Baseline Patterns of Application Behaviour

In this section, we report the main patterns on application behaviour. First, we present the overall patterns; then we explore heterogeneity by gender and socioeconomic background. We first deal with the number of applications, institutions, and fields. In the last part, we focus on the measures of selectivity.

4.1 Basic patterns

Table 2 reports in the upper panel the main statistics on application behaviour for the student population as a whole. On average, students submit 4.6 applications out of a possible maximum of 5.[12] So, despite the uncertainty related to admissions and the fact that it is in no way compulsory to accept an offer made by a university, applicants do not take full advantage of the possibility of submitting five applications. In terms of trends, Figure 4 (Panel A) shows the share of applicants who sent one to five applications via the main scheme route over time. A high and increasing number of applicants submit the maximum of five applications, while submitting four applications is the second most frequent behaviour, with around 10% of applicants doing so. The other categories are smaller but not negligible, hovering around 5% throughout the period.

Table 2: Baseline UCAS facts on applications, fields, institutions, and offers (2008–2021)

	Applications	Fields	Institutions	Applications in home region	Offers
Overall	4.59	1.63	4.59	1.84	3.66
Gender					
Male	4.61	1.65	4.57	1.79	3.80
Female	4.58	1.62	4.61	1.88	3.56
SES Status					
Low	4.51	1.64	4.53	2.21	3.50
Medium	4.59	1.64	4.58	1.83	3.68
High	4.69	1.61	4.68	1.45	3.83
Unknown	4.55	1.64	4.58	1.99	3.56
N	5,128,014	4,141,175	4,141,175	4,141,175	4,141,175

Notes: UCAS values pooled over 2008–2021. "Applications" uses the full applicant sample (1–5 applications). "Fields", "Institutions", "Applications in home region", and "Offers" are defined on the subset of applicants who submitted exactly five applications to ensure portfolios are comparable.

Focusing on students who submit five applications,[13] we find that in terms of number of fields, defined by 19 "subject areas" (e.g., in natural sciences there is a distinction between "Biological sciences" and "Physical sciences"), they include, on average, 1.6 distinct subject areas, while in terms of number of institutions, students apply to 4.6 different institutions on average. Panel B of Figure 4 shows the distribution of how many of the five applications are sent to home-region institutions.[14] Combined with the averages in Table 2, this confirms that applications to local institutions are common but far from universal: the vast majority of applicants send some applications outside their home region, and a non-

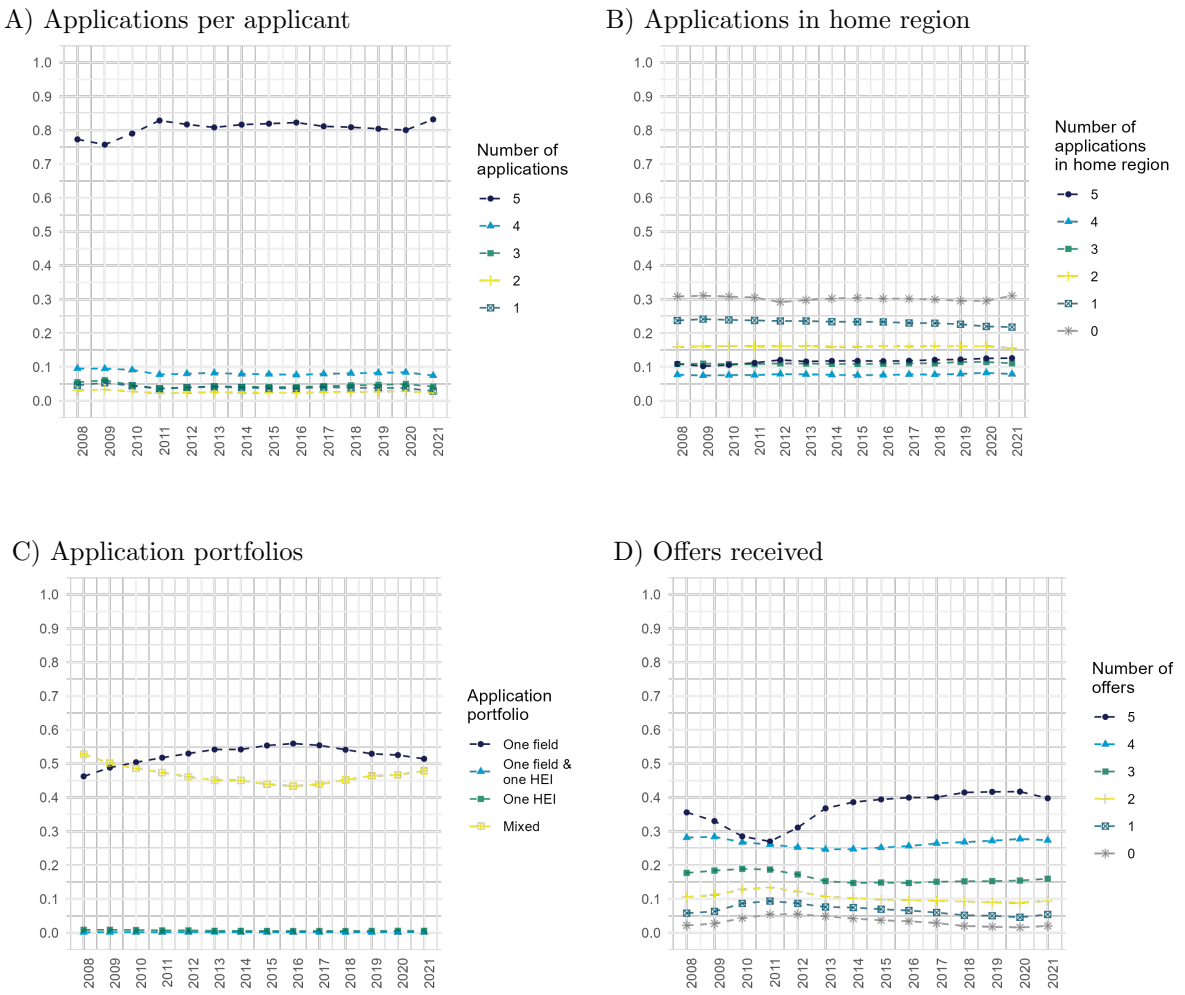
[12] Applications to medicine, dentistry and veterinary science are restricted to four. However, they represent only 4.2% of the overall sample, thus reducing the theoretical maximum only marginally, from 5 to 4.96.

[13] While the five-application subsample may introduce selection bias, this group is highly relevant for policy, as the majority of applicants submit five applications. Any variation by gender or SES is unlikely to significantly affect the conclusions, given the large share of the population in this subsample.

[14] Regions follow the UCAS classification Scotland, Northern Ireland, Wales, and for England: East Midlands, East of England, London, North East, North West, South East, South West, West Midlands, and Yorkshire and the Humber.

trivial share apply exclusively outside their home region. On the opposite side, there is a non-trivial share, above 10%, of applicants who send all five applications within their home-region.

Figure 4: Trends in Applications, Portfolios, and Offers 2008-21

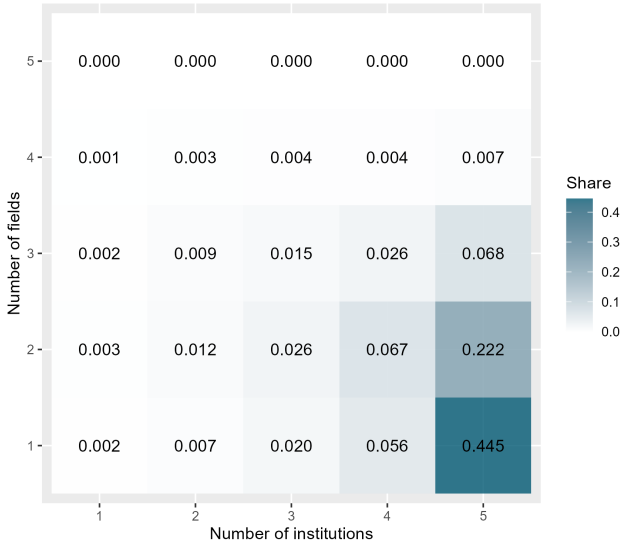


Notes: Panels A to D show time trends in applicant behaviours. Panel A depicts the number of applications per applicant (ranging from 1 to 5) for all main scheme applications, while Panels B to D are conditional on applicants having submitted five applications. Panel B shows applications to the home region, Panel C shows within-portfolio switching across fields and institutions, and Panel D shows the total offers received. For a distinction between conditional and unconditional offers over time please see Figure A3 in the Appendix. All values are presented as shares, with categories illustrated in different colors.

Figure 4 (Panel C) analyzes the joint choice of university and field, classifying students into four categories: those applying to one field and one institution, those applying to multiple fields within one institution (thus showing a strong institutional preference), those applying to multiple institutions within one field (thus showing a strong field preference), and the mixed category of those applying to more than one field in more than one institution. The first two categories are very marginal. The one-field, multiple-institutions category is instead the modal choice, accounting for between 45-55% of applicants, with a tendency for this group to increase until 2016 and decline in relative importance in recent years. The mixed group is also very relevant and in terms of trends is the mirror image of the previous category. To better understand this 'mixed' category, in Table A6 in the Appendix we provide the shares of the different application patterns. Regardless of the field definition, the most important category in the mixed group is the two-fields/five-institutions one, accounting for more than one-fifth of applicants. Thus, the strong field preference is quite evident.

The heat map (Figure 5) reveals a concentrated application pattern among undergraduate applicants, with the majority choosing to apply to five institutions. Applicants generally limit themselves to one or two fields of study, as shown by the high density in the bottom-right quadrant of the map.

Figure 5: Share of applicants across institutions and fields.



Notes: Heat map illustrating the distribution of applicants to undergraduate courses based on the number of institutions they apply to (x-axis) and the number of fields they apply in (y-axis). The density of applicants is highest in the bottom right quadrant, indicating that most applicants apply to four or five institutions, but typically focus on just one or two fields.

We also construct, for each applicant, a count of applications sent to institutions in the applicant's home region. This 'applications in home region' measure is defined on the five-application subsample and captures the extent to which applicants keep their portfolios geographically local. Applicants send on average 1.84 applications in their home region. There is little gender difference (1.79 for males versus 1.88 for females), but the SES gradient is stark: low-SES applicants submit 2.21 applications to institutions in their home region, compared to 1.45 for high-SES applicants. This is consistent with lower-SES students facing stronger geographical and financial constraints and therefore keeping more of their portfolio local.

Finally, we see that applicants receive on average 3.7 offers out of their 5 applications. Figure 4 (Panel D) shows trends in the number of offers. After a progressive decline up until 2011, we can see an increasing trend in the share of students receiving five offers, i.e. an offer for each of the five applications they submitted. The share of applicants receiving no offers is instead rather small, declining from a peak of 5% to just a few percentage points in recent years.

In the Appendix, we examine these choices in more detail. In particular, Figure A1, illustrates the share of applicants who apply to up to five different institutions (Panel A) and up to five different broad (Panel B) or narrow (Panel C) fields. Panel A shows that most students submit their applications to five distinct institutions. Yet, 25% of these applicants - who submit five applications - apply to fewer than five institutions. This means that they submit multiple applications to the same institution, even if the share of those making five applications to only one institution is minimal. Panels B and C show that there was a tendency up until 2016 for students to decrease the variety of fields they were applying to, but this has reversed in recent years. Indeed, when we consider the broad definition of field, we see that the share of those submitting all five applications within the same field increased approximately from 45% in 2008 to 55% in 2016, but then declined to 50%. Most multi-field applicants choose within two

broad fields, while there is a sizeable minority spanning three fields. When considering narrow fields in Panel C, the overall trends are rather similar, even if the levels are obviously different, as the distinction between fields is much more fine-grained.

Regarding offers, Figure A2 distinguishes between conditional (Panel A), unconditional (Panel B) offer in terms of numbers, while Figure A3 in terms of shares. We can see that unconditional offers are relatively rare, even if there has been a steady increase between 2013 and 2020 of the share of applicants receiving one or two (and to a lesser extent three) unconditional offers, followed by a drop in 2021.[15] We see instead that most applicants receive at least one conditional offer.

We next examine how these patterns vary by gender and socioeconomic background.

4.2 Heterogeneity by gender and socioeconomic background

First, we consider heterogeneity by gender. In terms of applications submitted, the numbers for males (4.61) and females (4.58) are very similar. Among those submitting five applications, the difference is small also in terms of number of fields chosen (1.65 for males vs. 1.62 for females) and in terms of number of institutions (4.57 for males vs. 4.61 for females). There is instead a more significant gender gap in terms of number of offers when submitting five applications, with male applicants receiving on average 3.80 offers and female ones 3.56. As we will see in the next section, females tend to apply to more selective courses.

Next, we consider differences between applicants with different socio-economic status (SES). We define low SES applicants as those whose household head is employed in a routine or semi-routine occupation, high SES applicants as those with at least one parent working in higher managerial or professional occupations. Medium SES are all others for whom socio-economic status is known. Given that there is a sizeable share of applicants for whom socio-economic status is unknown, we also include this category. There is a similar share of applicants classified in low (17%) and high (22%) SES status, while 45% are medium and for the rest the status is unknown (17%).

In the data, there is a gradient in terms of number of submitted applications: low SES applicants submit on average 4.51 applications, compared to 4.59 for medium SES and 4.69 for high SES applicants. Conditional on having submitted the maximum of five applications, there is not much of a difference in terms of the number of fields, while there is a gradient in the number of institutions students apply to: low SES apply on average to 4.53, medium SES to 4.58, and high SES to 4.68. Finally, there is a considerable difference in terms of offers received, with low SES students receiving on average 3.50 offers, compared to 3.68 for medium SES and 3.83 for high SES. In all instances, the "unknown" category is positioned between the low and medium SES categories. As we will see in the next section, low SES tend to apply to more selective courses than high SES applicants.

In the Appendix, we provide more detailed statistics in terms of heterogeneity by gender, SES, and indexes of deprivation (Polar4 and IMD). In particular, in Table A1, we explore heterogeneity in the number of applications submitted. We can see that females are slightly overrepresented among those submitting less than five applications, e.g., they are 56% of the overall population but 58% among those submitting 4 applications. The same is true for low SES applicants, while the high SES applicants are instead underrepresented among those submitting less than five applications. For those submitting five applications, we do the same for the number of institutions applied to (Table A2) and the number of

[15] The year 2021 is the first after the "Fair admission review" promoted by Universities UK (UUK), an organization putting together 141 vice-chancellors or principals of UK universities. UCAS writes regarding the end-of-cycle report for 2021: "The data on unconditional offers shows that universities have responded to recommendations in our Fair Admissions Review, aimed at building greater levels of transparency, fairness, and trust in the system, and worked hard to provide stability during the uncertainty caused by the pandemic." - see `https://www.ucas.com/corporate/news-and-key-documents/news/ucas-end-cycle-2021-strong-demand-uk-he-amidst-global-pandemic`

fields, distinguishing between a broad (Table A3) and a narrow (Table A4) definition. Female applicants are underrepresented among those who send their five applications to less than five different institutions and among those who choose more than one field, regardless of the definition. The same pattern is present for high SES applicants, while the opposite is true for those with low SES. The IMD indicator tends to decrease with the number of fields and increase with the number of institutions. This means that applicants from disadvantaged areas generally tend to choose a wider range of fields and fewer institutions. For the Polar4 indicator, patterns are more mixed.

Appendix Table A7 decomposes the five-application subsample by how many applications are sent to institutions in the applicant's home region. About 31% of applicants submit no home-region applications at all, 22% submit exactly one, and 14% submit all five to local institutions. Female applicants are slightly over-represented among those who submit many home-region applications. Low-SES students are under-represented among those with no home-region applications and over-represented among those with many, whereas high-SES students show the opposite pattern: they are more likely to submit zero local applications and less likely to submit portfolios concentrated in the home region. These patterns mirror the SES gradient in Table 2 and highlight the role of geography in shaping portfolio choices.

Finally, Table A8, does the same for socio-economic characteristics of students who receive 0 to 5 offers. Females and low SES applicants are strongly overrepresented among those receiving a low number of offers, while the opposite is true for high SES ones. Also, the IMD index is increasing in the number of offers received, indicating a less favourable outcome for applicants coming from disadvantaged areas.

To summarize, the analysis of basic patterns points to systematic variation in application and offers outcomes by gender and socio-economic background. After briefly discussing basic patterns related to the selectivity measure, in section 5, we will explore whether these differences are robust to controlling for several important co-variates.

4.3 Selectivity Profiles

Following the definition of selectivity we have provided in section 2.4, we see in Table 3 that the overall median is 0.31. This means that on average, the median application is to a place that does not provide an offer to 31% of applicants. The range of applications is quite wide, at 0.39. Indeed, on average the most selective application has a selectivity of 0.54, meaning that more than half of applicants do not receive an offer, while the least selective has a much lower level of 0.14. It is evident that female applicants apply to more selective places. Indeed, on average the median application for females has a selectivity of 0.34 vs 0.26 for males. The range is rather similar, but both most and least selective applications for females are on average more selective than the corresponding ones for males.

Table 3: Summary Statistics of Selectivity by Gender and Parental SES

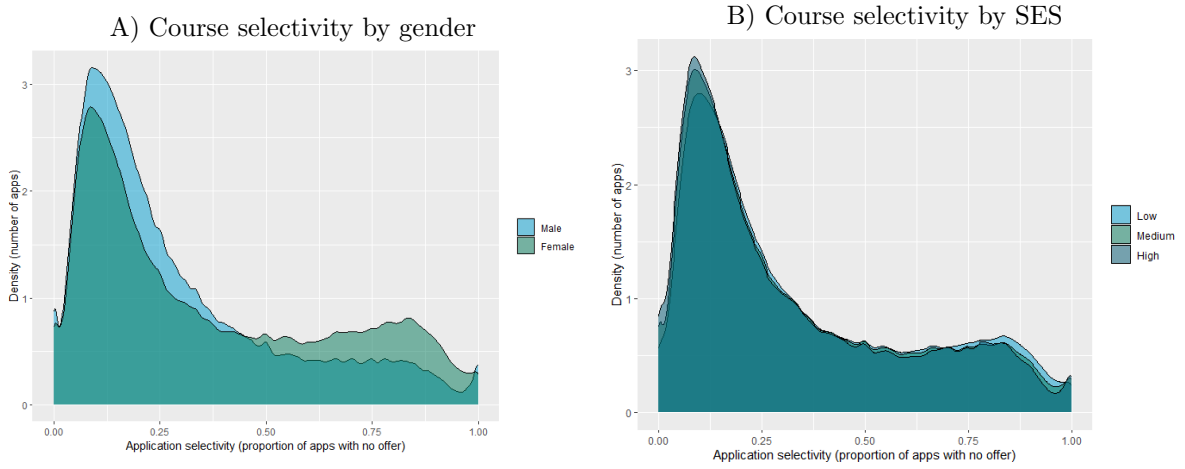
	Median	Range	Maximum	Minimum	N
Overall	0.31	0.39	0.54	0.14	4,141,175
Gender					
Male	0.26	0.39	0.50	0.12	1,836,293
Female	0.34	0.40	0.57	0.17	2,304,882
SES Status					
Low	0.32	0.38	0.54	0.16	685,117
Medium	0.30	0.39	0.53	0.14	1,843,826
High	0.30	0.41	0.54	0.13	919,266
Unknown	0.31	0.40	0.54	0.15	692,966

Notes: Summary statistics for median selectivity, range, maximum, and minimum selectivity by gender and parental SES status. N indicates the number of observations.

Panel A in Figure 6 shows the distribution of the selectivity of the applications by gender. The general shape of the two distribution is not too dissimilar. Both are right skewed, with a peak around 0.1. However, it is evident that females have a much larger mass in correspondence to high values, with the small exception of the very top, where males applicants are present in slightly larger number.

Instead, there is not much difference along the socio-economic status dimension. Low SES applicants apply to slightly more selective courses, with a smaller range in applications. Although the most selective applications are on average very similar, the least selective ones are actually lower for high SES applicants. As evident in Figure 6, the distributions are largely overlapping. Interestingly, as in the case of males, we can see an uptick at the very top of the distribution for high SES applicants.

Figure 6: Density plots of course selectivity by gender and SES status



Notes:

5 Conditional Evidence: Accounting for Ability and Selectivity

5.1 Empirical Framework

To deepen our understanding of these outcomes, we estimate the following specification:

$$Y_{it} = \alpha + X_{it}\beta + Q_{it} + A_{it}\lambda + \eta_{ct} + \varepsilon_{it}, \tag{2}$$

where Y_{it} is an application outcome variable (number of offers, applications, fields applied to etc.) for individual i in cycle t; X_{it} is a vector of individual demographic characteristics (POLAR, IMD, gender, ethnicity, socio-economic status - when not assessing heterogeneity along this dimension. See specifications for details); Q_{it} is a vector of A level qualification indicators (including grades), A_{it} is the number of A levels. We include country of domicile (England, Wales, Scotland and Northern Ireland) by time η_{ct}, fixed effects; ε_{it} are idiosyncratic errors.

Portfolio outcomes (fields, institutions, offers) are estimated on the five-application sample to keep portfolios comparable.

Selectivity To investigate strategic application behaviour, we also estimate specifications with our selectivity moments (see section 2.4) as the outcome variables. Our four selectivity outcomes include a) median selectivity θ_{it}, b) range of selectivity, c) highest selectivity and d) lowest selectivity. We run models in line with Equation 2, including the same set of control variables. However, we include deciles of median selectivity as a covariate for our range, highest and lowest selectivity outcome variables. This flexibly controls for average selectivity and means we can identify which groups submit more extreme (high or low selectivity) applications, conditional on a specific median selectivity.

5.2 Results on Applications, Subjects, Institutions, and Offers

We study four outcomes: applications (full sample) and, on the five-application subsample, fields, institutions, applications to institutions in the home region, and offers. Our results are presented in Table 5 and summarized in Figure 7 for gender and SES. Table 4 provides a guide to how coefficients should be interpreted for these portfolio and selectivity outcomes.

Table 4: How to read coefficients: portfolio and selectivity outcomes

Outcome Y	Unit/Scale	Positive coefficient on group indicator[a] means:
Applications	count (1–5)	More applications submitted
Subjects	count	Broader subject diversification
Institutions	count	Broader institution diversification
Offers	count	More offers received
Median selectivity	index in [0,1][b]	More selective overall portfolio
Range (max–min)	index in [0,1]	Wider vertical spread (more reach and/or more safety)
Max selectivity (most selective)	index in [0,1]	More reach at the top end (more ambitious top choice)
Min selectivity (least selective)	index in [0,1]	Less safety (even the safest choice is fairly selective)

[a] Group indicator is *Female* (vs. Male) or *Medium/High-SES* (vs. Low-SES), as specified in each table/figure.

[b] Selectivity is scaled so that higher values = *more selective* (lower offer rates). Median selectivity is the median across an applicant's five applications. For *Range/Max/Min* models we include deciles of median selectivity to condition on portfolio tier; the *Median selectivity* model excludes these by construction.

Notes: Applications use the full 1–5 sample. All other outcomes are estimated on the five-application subset.

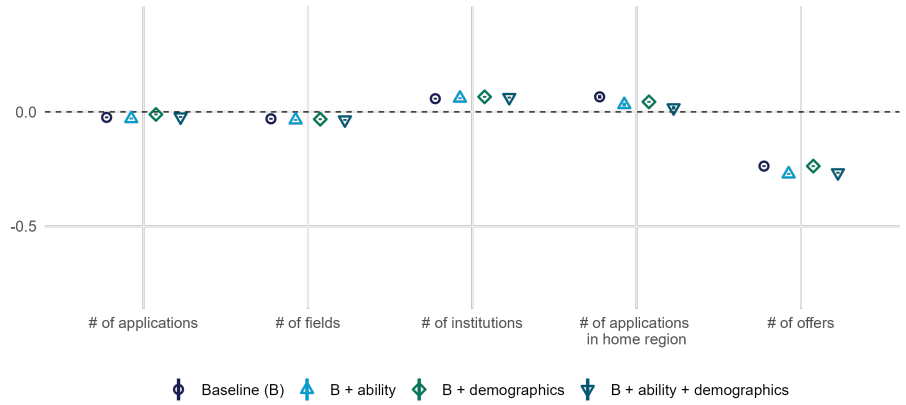
Gender Female applicants submit marginally fewer applications than males, choose fewer distinct fields but slightly more institutions, and receive fewer offers. In the case of applications, we see that relative to the unconditional regression, adding demographics reduces the gender gap, but does not eliminate it. Female applicants also direct more of their applications to institutions in their home region than males, even in the full specification with demographics and ability controls (column 3b), though the coefficient is very small 0.018. For the other outcomes, demographics do not affect the gender gap. When ability controls only are included the gender gap if anything widens slightly, while when both demographics and ability controls are included we see that gaps are either similar to the unconditional ones or slightly larger.

SES High-SES applicants submit more applications, diversify across more institutions, and receive more offers in the raw data. In the case of applications, we see that relative to the unconditional regression, adding demographics roughly halves the gap between high and medium and low SES. High-SES applicants also submit substantially fewer applications to institutions in their home region than low-SES applicants. In the SES-only specification (column 2a), medium-SES applicants submit around 0.39 fewer home-region applications and high-SES applicants around 0.76 fewer, relative to low-SES applicants. After controlling for demographics and ability (column 3b), these gaps shrink to about -0.14 and -0.30 respectively. Similarly, for the other outcomes, we find that demographics explain the SES gaps but do not fully eliminate them. When ability controls only are included the SES gap closes but does not evaporate, while when both demographics and ability controls are included we see that gaps are substantially reduced. For example, for offers while the unconditional gap between High and Low is 0.33 it shrinks to 0.044 when both demographics and ability are accounted for. This suggests that most of the raw SES advantage are compositional (ability and demographics).

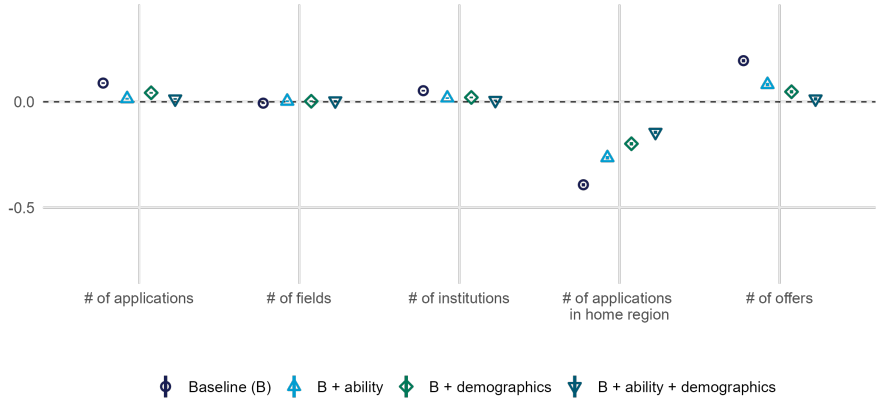
To summarize, gender gaps persist after rich controls, whereas SES gaps are largely explained by ability and other demographic controls, with small residuals. To further investigate whether these offer gaps reflect differential treatment within programs or compositional differences, we turn to application-level analysis where we can control for the specific program-institution combination.

Figure 7: Model specifications — variables of interest: Gender (female) and socio-economic background

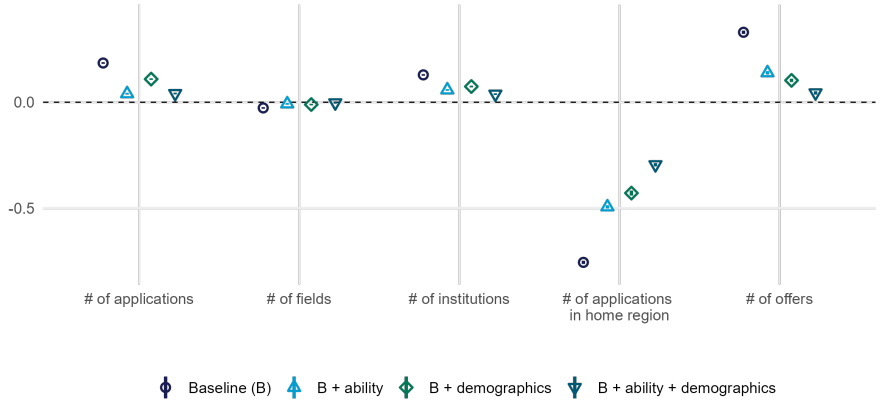
A) **Gender (female) coefficients** (99.99 % confidence intervals)



B) **Medium SES coefficients** (99.99 % confidence intervals)



C) **High SES coefficients** (99.99 % confidence intervals)



Notes: Point estimates and 99.99%-confidence intervals for variables "gender" (panel A), "medium SES" (panel B) and "high SES" (panel C) from regression results in Table 5. All specifications (including baseline) include country-by-time fixed effects. SES results are relative to low SES applicants. Low SES applicants are those whose household head is employed in a routine or semi-routine occupation, high SES applicants as those with at least one parent working in higher managerial or professional occupations. Medium SES are all others for whom socio-economic status is known. Demographic controls include age, minority ethnic status, POLAR quintiles, IMD quintiles, gender and SES status. Ability controls include UCAS tariff points, number of A levels and a set of dummy variables indicating whether students hold any of the ten most popular A levels and whether these were achieved with a grade of A or higher.

Table 5: Regression Results by Application Outcome

	Gender (1a)	Gender (1b)	SES (2a)	SES (2b)	Both (3a)	Both (3b)
Panel A: Applications						
Gender	-0.024***	-0.029***			-0.011***	-0.022***
Parental SES: Medium			0.089***	0.015***	0.043***	0.014***
Parental SES: High			0.185***	0.040***	0.109***	0.041***
Panel B: Fields						
Gender	-0.030***	-0.035***			-0.032***	-0.035***
Parental SES: Medium			-0.006***	0.003***	0.003***	0.005***
Parental SES: High			-0.027***	-0.008***	-0.011***	-0.003**
Panel C: Institutions						
Gender	0.058***	0.060***			0.066***	0.063***
Parental SES: Medium			0.053***	0.018***	0.021***	0.007***
Parental SES: High			0.129***	0.058***	0.074***	0.038***
Panel D: Applications to home region						
Gender	0.066***	0.032***			0.044***	0.018***
Parental SES: Medium			-0.392***	-0.264***	-0.198***	-0.144***
Parental SES: High			-0.756***	-0.494***	-0.429***	-0.295***
Panel E: Offers						
Gender	-0.237***	-0.271***			-0.237***	-0.266***
Parental SES: Medium			0.195***	0.082***	0.048***	0.014***
Parental SES: High			0.330***	0.139***	0.103***	0.044***
Controls						
Demographics	No	No	No	No	Yes	Yes
Ability	No	Yes	No	Yes	No	Yes

Notes: Robust standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All models include country of domicile by year fixed effects. $N = 5,128,014$ for panel A and $N = 4,141,175$ for panels B to E (sample only includes students who submitted five applications). Standard errors are always equal or below 0.002 and are excluded here for brevity. SES results are relative to low SES applicants. Low SES applicants are those whose household head is employed in a routine or semi-routine occupation, high SES applicants as those with at least one parent working in higher managerial or professional occupations. Medium SES are all others for whom socio-economic status is known. Demographic controls include age, minority ethnic status, POLAR quintiles, IMD quintiles, gender and SES status. Ability controls include UCAS tariff points, number of A levels and a set of dummy variables indicating whether students hold any of the ten most popular A levels and whether these were achieved with a grade of A or higher.

Offers conditional on subject, institution, and programme The gender gap in offers could arise due to: (i) females applying to more selective institutions within their chosen fields, (ii) females concentrating in subjects with inherently lower offer rates due to capacity constraints or higher competition, or (iii) differential evaluation or treatment of male and female applications by universities. The same potential mechanisms apply for high relative to low SES students. Understanding which mechanism drives the observed gaps has important implications for policy interventions aimed at reducing gender disparities in university admissions and widening participation to low SES students. To further probe the source of the gender and SES gaps in offers, Table 6 estimates application-level models with a Female indicator and SES indicators. The first two columns reproduce the patterns seen earlier - a female applicant is 4.7 pp less likely to receive an offer and a high SES student is 6.6 pp more likely to receive an offer than a low SES student. Like before, conditional on demographic and ability controls reported in column 3, the SES gap falls significantly - high SES students are only 0.9 pp more likely to receive an offer than low SES students. In contrast, females are still 5.3 pp less likely to receive an offer conditional on controls.

We sequentially add fixed effects: Institution FE (col. 4), Subject FE (col. 5), and Programme FE (course*institution; col. 6). In column 4, we find that adding Institution fixed effects leaves the female–male gap largely unchanged (col. 4). By contrast, Subject fixed effects absorb the gap (col. 5), which shrinks to 0.4 pp. When we include Programme (course–institution) fixed effects, the coefficient turns slightly positive (col. 6), indicating that, conditional on applying to the same programme, women are marginally more likely to receive an offer.

For the smaller SES gap, adding subject fixed effects (col. 5) leaves the gap unchanged whereas adding institution fixed effects (col.4) increases the gap. When we include programme fixed effects (col. 6), we find that high SES students are 2.3 pp more likely to receive an offer - similar to the female-male gap.

These positive effects (of being female and high SES) may be due to females and high SES applicants having slightly stronger unobservable dimensions of application quality (e.g., more polished personal statements, more enthusiastic reference letters, or better interview performance).

Together, these results imply that the relatively raw female shortfall in offers is driven by subject composition rather than by systematic differences in institution selectivity or within-programme treatment. We find no evidence of negative differential treatment against female applicants in offers. However, we find that conditional on applying to the same programme with similar characteristics and entry qualifications, females and high SES students are more likely to receive an offer.

Table 6: Regression Results Offers conditional on subject, institution, and programme

			Combined			
	Gender (1)	SES (2)	Controls (3)	Inst. FE (4)	Subject FE (5)	Course FE (6)
Female	-0.047***		-0.053***	-0.047***	0.004***	0.023***
Parental SES: Medium		0.039***	0.003***	0.009***	0.004***	0.011***
Parental SES: High		0.066***	0.009***	0.026***	0.008***	0.023***
N	20,705,875	20,705,875	20,705,875	20,705,875	20,705,875	20,705,875
R^2	0.020	0.020	0.058	0.168	0.162	0.411
Controls			✓	✓	✓	✓
FE: Institution				✓		
FE: Subject					✓	
FE: Course						✓

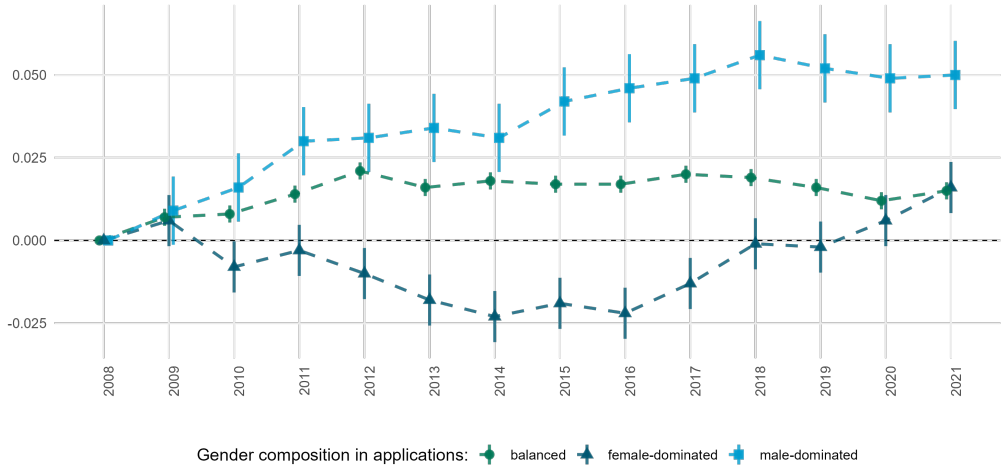
Notes: Standard errors clustered at the individual level. Standard errors are always below 0.0005 and are excluded here for brevity. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. All models include country of domicile by year fixed effects. Application-level linear probability models. Outcome $Y = 1$ if an application receives an offer (conditional or unconditional). *Female*= 1 for female applicants (male baseline). Controls include all demographic and ability controls.

Figure 8 illustrates how the marginal effect of being female on the number of offers evolves over time separately for male-dominated, balanced, and female-dominated fields. Interestingly, we see an upward trend in the male-dominated fields, indicating perhaps that is recent years Universities actively attempt to rebalance gender in those fields.

5.3 Application Strategy: Range, Safety, and Reach

We examine application strategies using four selectivity moments: median selectivity, range (max–min), maximum selectivity (most selective course), and minimum selectivity (least selective course). Results on application strategy are presented in Table 7 and summarized in Figure 9 for gender and SES.

Figure 8: Marginal effects of gender on the number of offers received over time



Notes: The figure illustrates the development of marginal effects of the interaction between gender and application cycle over time for three models: male-dominated fields, balanced fields, and female-dominated fields. These categories are defined according to the terciles of the gender composition in applications in 2008. The estimates represent the coefficients and 99 %-confidence intervals from regressions of the number of offers on gender, including country of domicile by year fixed effects and all demographic and ability controls.

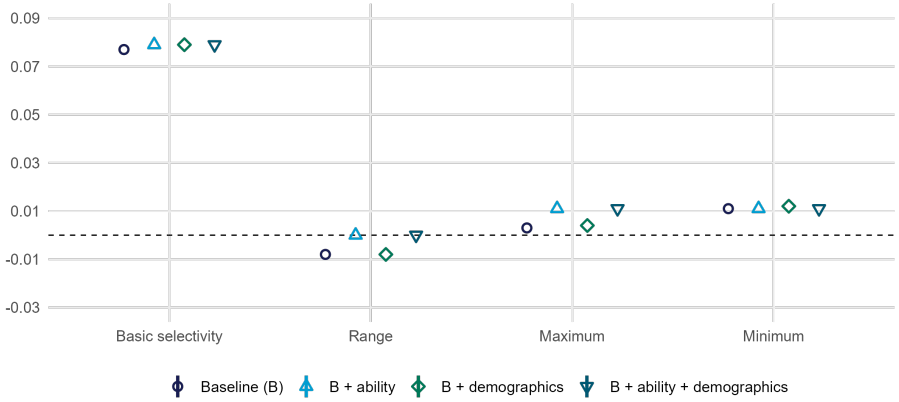
Gender Starting with patterns on unconditional strategy, controlling only for country of domicile by year fixed effects, we see that female applicants submit more selective applications on average, with a slightly smaller spread, driven by them being more ambitious for both their most selective and least selective courses. These gender differences in portfolio selectivity are not attenuated by controls for demographic characteristics or A-level attainment. The higher median selectivity and narrower range indicate that female applicants are not simply adding more "safe" options, instead, they shift the entire portfolio toward more selective courses. This pattern is consistent with gender differences in preferences over fields and in perceived admission risk, rather than systematic differences in safety behaviour per se.

SES Regarding SES, we find again different patterns. Medium and High SES groups submit lower selective applications on average, but this is explained by demographics and ability. Once those are accounted for, these groups submit higher selective applications on average, with a broader range, due to both higher ambitious and lower safe applications, although these differences are quantitatively small (e.g. the high-SES coefficient on median selectivity is around 0.001 in the fully controlled specification).

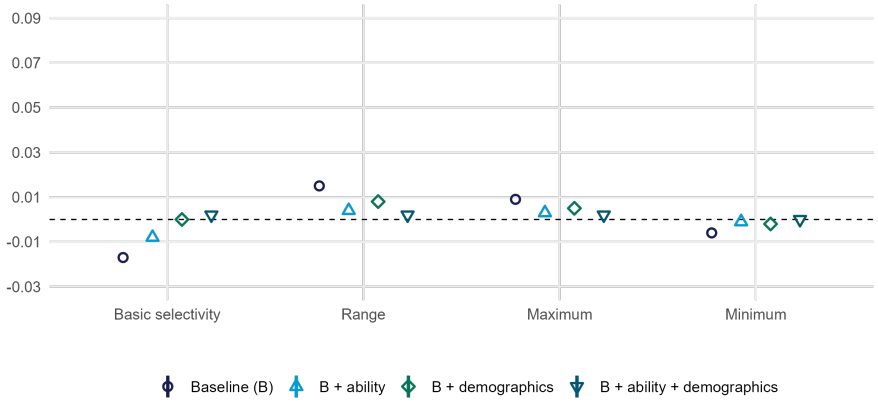
To summarize, in the case of gender, demographics, and ability controls do not explain much of the gap, whereas for SES they do.

Figure 9: Selectivity patterns — variables of interest: Gender (female) and socio-economic background

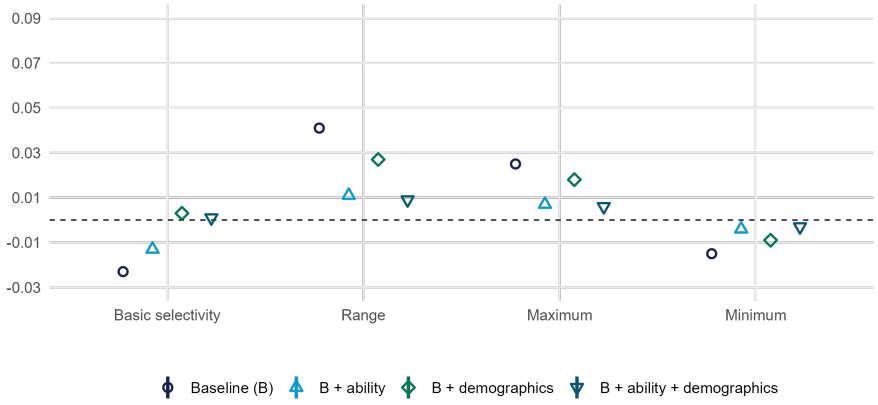
A) **Gender (female) coefficients** (99.99 % confidence intervals)



B) **Medium SES coefficients** (99.99 % confidence intervals)



C) **High SES coefficients** (99.99 % confidence intervals)



Notes: Point estimates and 99.99%-confidence intervals for variables "gender" (panel A), "medium SES" (panel B) and "high SES" (panel C) from regression results in Table 7. All specifications (including baseline) include country-by-time fixed effects. SES results are relative to low SES applicants. Low SES applicants are those whose household head is employed in a routine or semi-routine occupation, high SES applicants as those with at least one parent working in higher managerial or professional occupations. Medium SES are all others for whom socio-economic status is known. Demographic controls include age, minority ethnic status, POLAR quintiles, IMD quintiles, gender and SES status. Ability controls include UCAS tariff points, number of A levels and a set of dummy variables indicating whether students hold any of the ten most popular A levels and whether these were achieved with a grade of A or higher.

Table 7: Regression Results by Selectivity Outcome

	Gender (1a)	Gender (1b)	SES (2a)	SES (2b)	Both (3a)	Both (3b)
Panel A: Median selectivity						
Gender	0.077***	0.079***			0.079***	0.079***
Parental SES: Medium			-0.017***	-0.008***	0.000	0.002***
Parental SES: High			-0.023***	-0.013***	0.003***	0.001***
Panel B: Range of selectivity						
Gender	-0.008***	-0.000			-0.008***	-0.000
Parental SES: Medium			0.015***	0.004***	0.008***	0.002***
Parental SES: High			0.041***	0.011***	0.027***	0.009***
Panel C: Maximum selectivity						
Gender	0.003***	0.011***			0.004***	0.011***
Parental SES: Medium			0.009***	0.003***	0.005***	0.002***
Parental SES: High			0.025***	0.007***	0.018***	0.006***
Panel D: Minimum selectivity						
Gender	0.011***	0.011***			0.012***	0.011***
Parental SES: Medium			-0.006***	-0.001***	-0.002***	-0.000
Parental SES: High			-0.015***	-0.004***	-0.009***	-0.003***
Controls						
Demographics	No	No	No	No	Yes	Yes
Ability controls	No	Yes	No	Yes	No	Yes

Notes: All models include country of domicile by year fixed effects. Panels B–D flexibly control for median selectivity using deciles. N = 4,141,175 for panels A to D (sample only includes students who submitted five applications). Standard errors are always below 0.0005 and are excluded here for brevity. SES results are relative to low SES applicants. Low SES applicants are those whose household head is employed in a routine or semi-routine occupation, high SES applicants as those with at least one parent working in higher managerial or professional occupations. Medium SES are all others for whom socio-economic status is known. Demographic controls include age, minority ethnic status, POLAR quintiles, IMD quintiles, gender and SES status. Ability controls include UCAS tariff points, number of A levels and a set of dummy variables indicating whether students hold any of the ten most popular A levels and whether these were achieved with a grade of A or higher.

6 External Expectations

6.1 Expert Expectations

We next use expert expectations as a benchmark to compare predicted versus actual application behaviour. Table 8 compares UCAS averages for applications, fields, and offers with the mean and median predictions of education economists and UK academics, while Table 9 reports the distribution of their beliefs about gender gaps in applications and offers.

The mean number of applications, 4.59, exceeds the overall expert prediction of 3.88, with a stronger underprediction by economists in the field of education - who expect 3.55 - compared to UK academics - who expect 4.19, a number rather close to the actual one. Looking at the median prediction, this is 3 for education economists and 4 for UK academics. As we saw earlier, the number of applications has increased over time, so the underprediction by experts is not due to them giving excessive weight to recent years when evaluating the phenomenon over the 2008-2021 time period.

Experts also substantially underpredict the number of offers received by applicants who submit the maximum number of applications. The actual figure, that includes both conditional and unconditional

offers, is 3.66, but educational economists expected only 2.38 (median: 2), while UK academics considered 2.98 (median: 3), a smaller but still considerable difference.

On gender, the most common response (45% of experts) was that females submit more applications; 19% thought they submit fewer, and 17% expected no gender difference. In terms of offers the picture is more balanced, with 38% expecting females to receive more offers, 36% expecting no differences and only 11% expecting females to receive less offers. Also in this case education economists are more prone to have incorrect predictions than UK academics, with 51% and 52% expecting a positive female gap in applications and offers, despite the gap being in reality negative.

In the case of SES, experts are mostly correct in their predictions, with 79% correctly predicting that low SES students submit fewer applications, and 77% that they received fewer offers after sending five applications. In this case, education economists are more likely to be correct than UK academics.

Table 8: Comparison of UCAS data with expert predictions

		Survey data					
	UCAS data	**Pooled**		**Education Econ.**		**UK Academics**	
		Mean	Median	Mean	Median	Mean	Median
no. of appl.	4.59	3.88	4.00	3.55	3.00	4.19	4.00
no. of fields	1.63	2.05	2.00	2.22	2.00	1.89	2.00
no. of inst.	4.59	3.81	4.00	3.49	3.00	4.12	4.00
no. of offers	3.66	2.69	3.00	2.38	2.00	2.98	3.00
No. of obs.	4,141,175	124	124	61	61	63	63

Note: UCAS values are pooled over 2008-2021. The mean number of applications is based on N = 5,128,014 applicants who sent 1 to 5 applications. The mean values for no. of fields, institutions, and offers are calculated for applicants who sent five applications via the main scheme, i.e., on N = 4,141,175 applicants. Numbers for UCAS data are based on Table 2.

Table 9: Comparison of UCAS data with expert predictions by gender

	UCAS data	Survey data		
		Pooled	Education Economists	Academics in the UK
Number of applications:				
Average value for male applicants:	4.61			
Average value for female applicants:	4.58			
Expert predictions:				
Females sent fewer applications		0.19	0.28	0.11
Females sent more applications		0.45	0.51	0.40
No gender difference		0.17	0.11	0.22
Self-reported lack of knowledge		0.19	0.10	0.27
No. of observations	5,128,014	124	61	63
Number of offers received:				
Average value for male applicants:	3.80			
Average value for female applicants	3.56			
Expert predictions:				
Females receive less offers		0.11	0.08	0.14
Females receive more offers		0.38	0.52	0.24
No gender difference		0.35	0.34	0.37
Self-reported lack of knowledge		0.15	0.05	0.25
No. of observations	4,141,175	124	61	63

Note: Mean values pooled over the years 2008 to 2021. For the number of offers received, we condition on having sent five applications. Thus, the number of observations is lower. *Self-reported lack of knowledge* relates to the fourth option "I have no clue". Numbers may not add to 1 due to rounding. Numbers for UCAS data are based on Table 2.

6.2 ChatGPT Predictions

As a complement to the expert survey presented earlier, we use a large language model (ChatGPT) as an alternative benchmark to assess actual student application behaviour. Comparing the model's predictions to the observed data helps us identify which patterns reflect widely available information and which remain unexpected, even to a well-informed AI whose responses were informed by external data sources, e.g., the relevant UCAS End-of-Cycle Reports, approximating what a well-informed observer might predict.

Table 10 presents observed statistics from our dataset alongside estimates generated by ChatGPT o3 across different temperature settings.[16] Quantitatively, ChatGPT slightly underestimates several true values across all temperature settings, such as the number of applications and fields. The number of offers shows the greatest variation and ranges from a low of 2.7 at temperature 0.3 to a high of 3.8 at 0.1, compared to the true value of 3.7.

Regarding the raw differences in applications and offers between different groups, ChatGPT predicts varying trends for female applicants, including expecting them to send/receive more or fewer applications/offers than males, even though the actual data shows no difference in the number of applications sent by females but that they receive fewer offers. Regarding the raw differences by SES, ChatGPT correctly identifies that low-SES applicants tend to submit fewer applications and receive fewer offers,

[16]Temperature in language models like ChatGPT controls the randomness of the output with lower settings (e.g., 0.1) leading to more deterministic and consistent answers and higher temperatures (e.g., 0.9) producing more varied and potentially creative results.

aligning with the data and outperforming expert predictions on this dimension.

When controlling for socio-demographic factors, ability, and selectivity, ChatGPT's predictions diverge from actual patterns. For instance, it predicts that females will submit more applications than males, despite no difference in the data, and its expectations for the number of offers females would receive are inconsistent. Similarly, ChatGPT anticipates lower application numbers and offers for low-SES applicants, while the data shows no such disparity when accounting for selectivity and academic ability.

Table 10: True values compared to results provided by ChatGPT

	True value	**ChatGPT o3 (temperature)**				
		0.1	**0.3**	**0.5**	**0.7**	**0.9**
Average values:						
No. of applications	4.6	4.1	4.0	4.0	4.2	4.1
No. of fields	1.6	1.3	1.2	1.3	1.3	1.4
No. of institutions	4.6	4.3	4.9	4.2	4.3	4.3
No. of offers	3.7	3.8	2.7	3.4	3.3	3.6
Raw differences by gender and SES:						
Females send ... applications than males	same	same	same	fewer	same	more
Females receive ... offers than males	fewer	same	fewer	more	fewer	fewer
Low SES send ... applications than high SES	fewer	fewer	fewer	fewer	fewer	fewer
Low SES receive ... offers than high SES	fewer	fewer	fewer	fewer	fewer	fewer
Conditional differences by gender and SES:						
Females send ... applications than males	same	more	more	more	more	more
Females receive ... offers than males	fewer	fewer	more	same	more	more
Low SES send ... applications than high SES	same	fewer	fewer	fewer	fewer	fewer
Low SES receive ... offers than high SES	same	fewer	fewer	fewer	fewer	fewer

Note: Results are based on OpenAI o3 with the above stated temperature settings of 0.1, 0.3, 0.5, 0.7 (default), and 0.9. The o-series is trained to 'think' longer before responding. All responses have been backed by additional data, often extracted from UCAS' End-of-Cycle Reports. Full protocols of the interactions are available on request.

7 Conclusion

Using administrative records on all UK undergraduate applications from 2008–2021, this paper provides new evidence on how applicants construct university portfolios in a setting where subject and institution must be chosen jointly. We believe that this study provides a rich empirical baseline for future work and has relevant policy implications.

We document how applicants predominantly organise their choices around subjects rather than institutions, with roughly half of five-application portfolios consisting of a single subject area applied to across multiple universities. This pattern suggests that, even in a formally joint-choice environment, applicants structure decisions hierarchically and place primary weight on the field of study. This has implications, for instance, in terms of guidance programmes for university studies. Many such activities focus on a specific institution, e.g., on-campus open days. Universities clearly have a strong incentive to organize them, but our results suggest that applicants may find even more useful guidance activities focused on subjects. The UK initiative *Discover Economics*,[17] for instance, allows students to get a glimpse at what studying economics entails, without focusing on a specific university. Given students' behaviour, high

[17] https://www.discovereconomics.co.uk/

schools may want to prioritize such discipline-specific activities in their offer. Academies or scientific societies, like the RES in the UK or the AEA in the US, could provide support for such initiatives, as they are more focused on a specific discipline than the typical university with a broad portfolio of courses.

We also document sizable socioeconomic gradients in portfolio breadth, geographic scope, and offer rates. High-SES applicants submit more applications, apply to more institutions and fewer in their home region, and receive more offers. These differences, however, attenuate sharply once we control for prior attainment, indicating that much of the SES advantage reflects earlier inequalities in preparation and achievement rather than disparities in admissions decisions themselves. In the UK and elsewhere, universities are sometimes accused of elitism or classism.[18] Our results suggest that the concern about disparities in access is real, but the roots are not at the application or admission level, thus highlighting the importance of earlier interventions to address disparities of opportunities. The SES gradient in local applications is the only one that maintains a relevant magnitude after controlling for previous achievement and demographics. This underlines the importance of travel costs, information and housing constraints, and calls for specific policies to address them.

Finally, we show that gender differences arise in both application behaviour and admissions outcomes, even after controlling for previous achievements or demographics. Because they apply to more selective courses, women receive fewer offers than men despite submitting a similar number of applications. Application-level analyses reveal that these gaps disappear—and even reverse slightly—once we condition on programme-level fixed effects, implying that the observed differences are driven by the distribution of subjects to which women apply, rather than by differential treatment within programmes. This evidence runs counter to experts' expectations and thus can contribute to better anchoring the public debate on facts.

As mentioned in the introduction, education shapes both individual and societal outcomes. Debates on such an important topic should be grounded as much as possible on a shared understanding. We hope that this paper contributes to this with regard to the university application process.

References

Altmejd, A., Barrios-Fernández, A., Drlje, M., Goodman, J., Hurwitz, M., Kovac, D., ... Smith, J. (2021). O brother, where start thou? Sibling spillovers on college and major choice in four countries. *The Quarterly Journal of Economics*, *136*(3), 1831–1886.

Altonji, J. G., Arcidiacono, P., & Maurel, A. (2016). The analysis of field choice in college and graduate school: Determinants and wage effects. In *Handbook of the economics of education* (Vol. 5, pp. 305–396). Elsevier.

Arcidiacono, P., Hotz, V. J., & Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, *166*(1), 3–16.

Barrios-Fernández, A. (2022). Neighbors' effects on university enrollment. *American Economic Journal: Applied Economics*, *14*(3), 30–60.

Beffy, M., Fougere, D., & Maurel, A. (2012). Choosing the field of study in postsecondary education: Do expected earnings matter? *Review of Economics and Statistics*, *94*(1), 334–347.

Belfield, C., Boneva, T., Rauh, C., & Shaw, J. (2020). What drives enrolment gaps in further education? the role of beliefs in sequential schooling decisions. *Economica*, *87*(346), 490–529.

Blom, E., Cadena, B. C., & Keys, B. J. (2021). Investment over the business cycle: Insights from college major choice. *Journal of Labor Economics*, *39*(4), 1043–1082.

[18] See, for instance, https://www.independent.co.uk/news/uk/gap-england-department-for-education-government-data-b2634966.html or https://www.independent.co.uk/news/uk/ucas-university-of-exeter-alevel-england-wales-b2591678.html or https://www.theguardian.com/uk-news/2024/dec/26/scottish-students-at-edinburgh-university-start-support-group-to-counter-alienation or https://www.theguardian.com/education/2024/sep/12/radical-action-more-state-school-pupils-oxbridge-oxford-cambridge

Bordon, P., & Fu, C. (2015). College-major choice to college-then-major choice. *The Review of Economic Studies*, *82*(4), 1247–1288.

Britton, J., Van Der Erve, L., Belfield, C., Vignoles, A., Dickson, M., Zhu, Y., ... Buscha, F. (2022). How much does degree choice matter? *Labour Economics*, *79*, 102268.

Broecke, S. (2015). University rankings: do they matter in the uk? *Education Economics*, *23*(2), 137–161.

Calsamiglia, C., Fu, C., & Güell, M. (2020). Structural estimation of a model of school choices: The boston mechanism versus its alternatives. *Journal of Political Economy*, *128*(2), 642-680. doi: 10.1086/704573

Campbell, S., Macmillan, L., Murphy, R., & Wyness, G. (2022). Matching in the dark? Inequalities in student to degree match. *Journal of Labor Economics*, *40*(4), 807–850.

Chevalier, A., & Jia, X. (2016). Subject-specific league tables and students' application decisions. *The Manchester School*, *84*(5), 600–620.

Declercq, K., & Verboven, F. (2015). Socio-economic status and enrollment in higher education: do costs matter? *Education Economics*, *23*(5), 532–556. doi: 10.1080/09645292.2015.1047822

Delaney, J. M., & Devereux, P. J. (2019). Understanding gender differences in stem: Evidence from college applications. *Economics of Education Review*, *72*, 219–238.

Delavande, A., & Zafar, B. (2019). University choice: The role of expected earnings, nonpecuniary outcomes, and financial constraints. *Journal of Political Economy*, *127*(5), 2343–2393.

DellaVigna, S., Pope, D., & Vivalt, E. (2019). Predict science to improve science. *Science*, *366*(6464), 428-429. doi: 10.1126/science.aaz1704

Gibbons, S., Neumayer, E., & Perkins, R. (2015). Student satisfaction, league tables and university applications: Evidence from britain. *Economics of Education Review*, *48*, 148–164.

Gibbons, S., & Vignoles, A. (2012). Geography, choice and participation in higher education in England. *Regional Science and Urban Economics*, *42*(1-2), 98–113. doi: 10.1016/j.regsciurbeco.2011.07.004

Goldin, C., & Katz, L. F. (2009). *The race between education and technology.* harvard university press.

Hoxby, C. M., & Avery, C. (2012). *The missing "one-offs": The hidden supply of high-achieving, low income students* (Tech. Rep.). National Bureau of Economic Research.

Lovenheim, M., & Smith, J. (2023). Returns to different postsecondary investments: Institution type, academic programs, and credentials. In *Handbook of the economics of education* (Vol. 6, pp. 187–318). Elsevier.

Patnaik, A., Wiswall, M., & Zafar, B. (2021). College majors 1. *The Routledge Handbook of the Economics of Education*, 415–457.

Pigini, C., & Staffolani, S. (2015). Beyond participation: do the cost and quality of higher education shape the enrollment composition? the case of italy. *Higher Education*, *71*, 119-142. doi: 10.1007/s10734-015-9892-8

Schmidt, A., Ortúzar, J. d. D., & Paredes, R. D. (2019). Heterogeneity and college choice: Latent class modelling for improved policy making. *Journal of Choice Modelling*, *33*(September). doi: 10.1016/j.jocm.2019.100185

Sloane, C. M., Hurst, E. G., & Black, D. A. (2021). College majors, occupations, and the gender wage gap. *Journal of Economic Perspectives*, *35*(4), 223–248.

Stinebrickner, R., & Stinebrickner, T. R. (2014). A major in science? Initial beliefs and final outcomes for college major and dropout. *Review of Economic Studies*, *81*(1), 426–472.

UCAS. (2022). *Universities and Colleges Admissions Service.* ONS SRS Metadata Catalogue. Dataset. (Released 02 August 2022) doi: 10.57906/qwbe-x221

Wiswall, M., & Zafar, B. (2015). Determinants of college major choice: Identification using an information experiment. *The Review of Economic Studies*, *82*(2), 791–824.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources*, *48*(3), 545–595.

A Additional Tables and Figures

Table A1: Descriptive statistics of applicants by number of applications sent

	Number of applications via main scheme					
	1	2	3	4	5	Overall
Share female	0.58	0.58	0.58	0.58	0.56	0.56
Polar4	3.19	3.25	3.24	3.37	3.53	3.48
IMD	3.14	3.20	3.19	3.29	3.30	3.29
Parental SES:						
Low SES	0.23	0.21	0.2	0.18	0.17	0.17
Medium SES	0.43	0.45	0.46	0.45	0.45	0.45
High SES	0.14	0.15	0.16	0.20	0.22	0.21
Unknown SES	0.20	0.18	0.18	0.17	0.17	0.17
N	202213	132355	229622	422649	4141175	5128014
Shares	0.04	0.03	0.04	0.08	0.81	1.00

Note: Mean values pooled over the years 2008 to 2021. Polar4 classifies small areas across the UK into quintiles ranging from quintile 1 (areas with the lowest young participation rates, considered as the most disadvantaged) to quintile 5 (highest young participation rates, considered most advantaged). The Index of Multiple Deprivation (IMD) identifies small area concentrations of multiple deprivation across all of the UK and thereby provides a relative measure of deprivation among small areas ranging from quintile 1 (highest deprivation) to quintile 5 (lowest deprivation). Parental SES is based on the following categories: High SES includes higher managerial and professional occupations, medium SES includes lower managerial and professional occupations, intermediate occupations, small employers and own account workers as well as lower supervisory and technical occupations. Low SES includes semi-routine and routine occupations.

Table A2: Descriptive statistics of applicants by number of institutions applied to

	Number of institutions a student applied to					
	1	2	3	4	5	Overall
Share female	0.46	0.49	0.54	0.56	0.56	0.56
Polar4	3.39	3.45	3.39	3.47	3.55	3.53
IMD	2.95	3.07	3.12	3.27	3.34	3.30
Parental SES:						
Low SES	0.22	0.20	0.19	0.17	0.16	0.17
Medium SES	0.47	0.48	0.46	0.45	0.44	0.45
High SES	0.12	0.15	0.17	0.21	0.23	0.22
Unknown	0.18	0.17	0.18	0.17	0.17	0.17
N	32457	129163	268893	633883	3076779	4141175
Shares	0.01	0.03	0.06	0.15	0.74	1.00

Note: Mean values pooled over the years 2008 to 2021. For definition of Polar4, IMD, and Parental SES, please refer to Table A1.

Table A3: Descriptive statistics of applicants by number of broad fields applied to

	Number of fields a student applied to					
	1	2	3	4	5	Overall
Share female	0.56	0.55	0.54	0.53	0.51	0.56
Polar4	3.52	3.54	3.51	3.52	3.55	3.53
IMD	3.32	3.3	3.24	3.17	3.10	3.30
Parental SES:						
Low SES	0.17	0.16	0.17	0.18	0.18	0.17
Medium SES	0.44	0.44	0.45	0.46	0.48	0.45
High SES	0.23	0.23	0.2	0.19	0.17	0.22
Unknown	0.17	0.17	0.17	0.17	0.17	0.17
N	2193923	1366274	500413	75309	5256	4141175
Shares	0.53	0.33	0.12	0.02	0.00	1.00

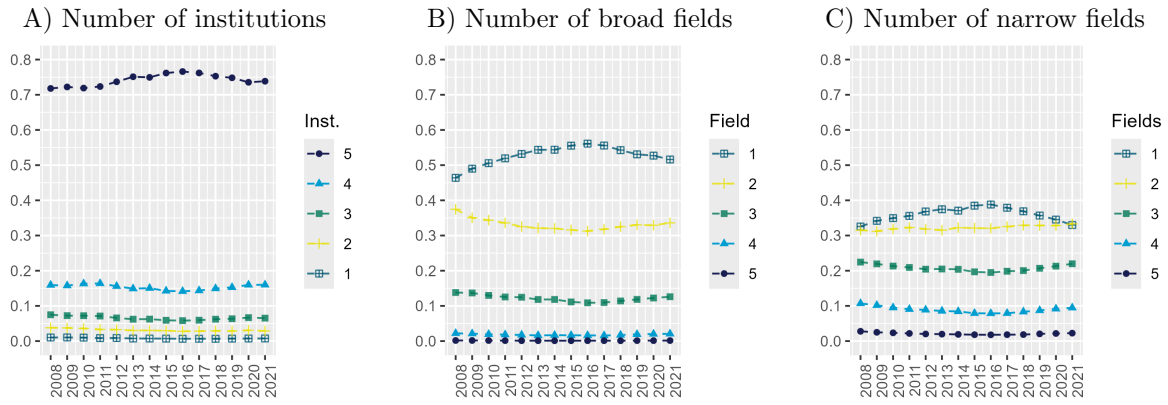
Note: Mean values pooled over the years 2008 to 2021. Field defined as first digit of JACS3 code includes combined degrees (19 different fields). For definition of Polar4, IMD, and Parental SES, please refer to Table A1.

Table A4: Descriptive statistics of applicants by number of narrow fields applied to

	Number of fields a student applied to					
	1	2	3	4	5	Overall
Share female	0.59	0.54	0.53	0.53	0.52	0.56
Polar4	3.54	3.53	3.50	3.50	3.52	3.53
IMD	3.36	3.31	3.26	3.19	3.14	3.30
Parental SES:						
Low SES	0.16	0.16	0.17	0.18	0.19	0.17
Medium SES	0.44	0.44	0.45	0.45	0.46	0.45
High SES	0.23	0.23	0.21	0.19	0.18	0.22
Unknown	0.16	0.17	0.17	0.18	0.18	0.17
N	1495124	1335664	858220	366004	86163	4141175
Shares	0.36	0.32	0.21	0.09	0.02	1.00

Note: Mean values pooled over the years 2008 to 2021. Field defined as exact JACS3 code (192 different fields). For definition of Polar4, IMD, and Parental SES, please refer to Table A1.

Figure A1: Application patterns

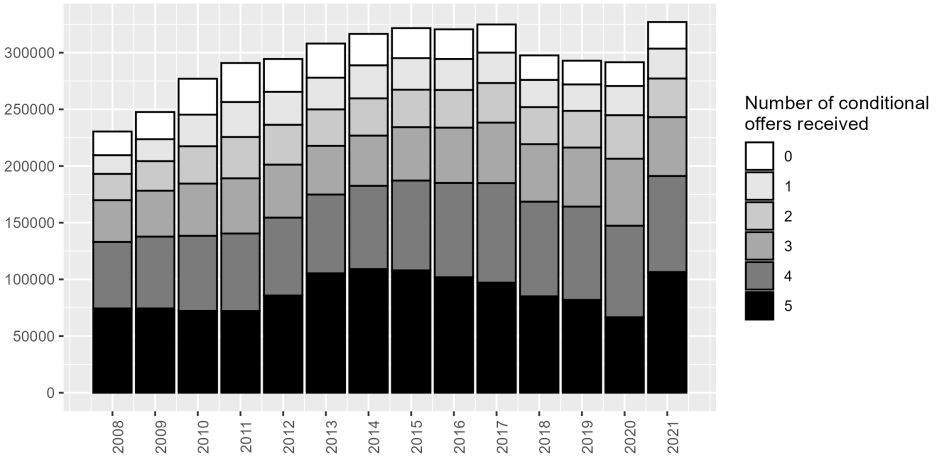


Notes: Panel A: Share of applicants who submit applications to one to five different institutions. Panel B: Share of applicants who submit applications to one to five different fields. Definition of field is based on letter of JACS3 (first digit). We distinguish between 19 different fields. Combined degrees count as independent fields. Panel C: as in B but definition of fields is based on two digits of JACS3. We distinguish between 192 different fields in Panel C. For a complete list of fields, see Table B1 in Appendix B. Combined degrees are included in the figure.

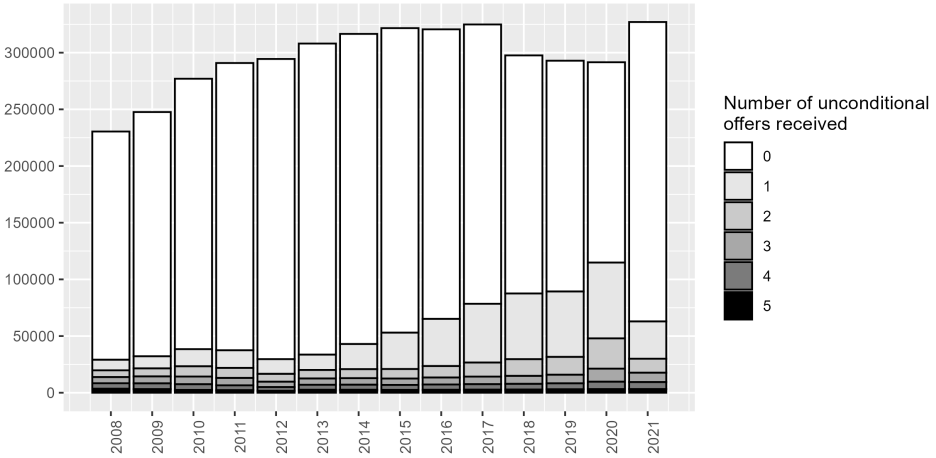
In Table A5, we explore heterogeneity between these four categories, when considering broad fields. Looking at the two most relevant ones, i.e., One field and Mixed, there are no significant differences regarding SES or age, and also the underrepresentation of females and overrepresentation of minorities in the mixed category are not quantitatively very large.

Figure A2: Number of offers

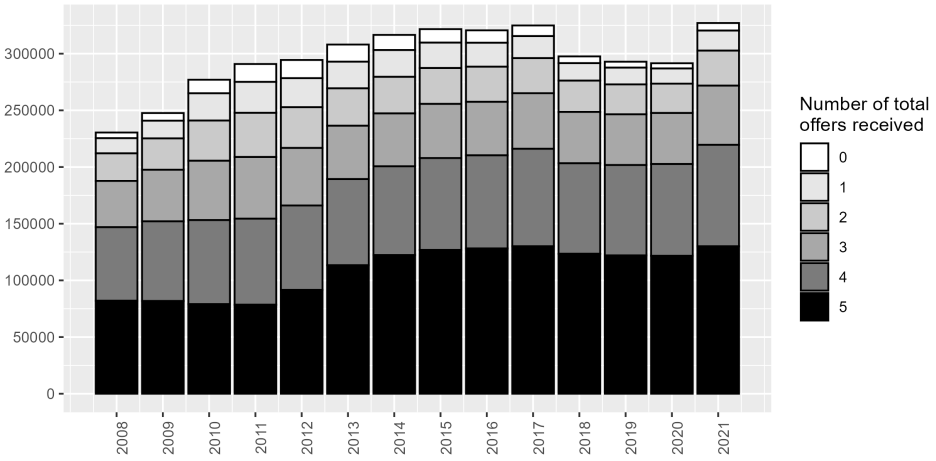
A) Conditional offers



B) Unconditional offers



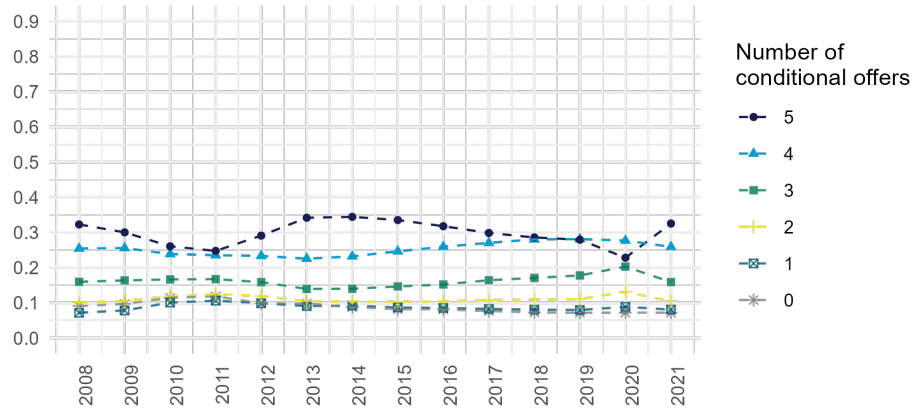
C) Total offers



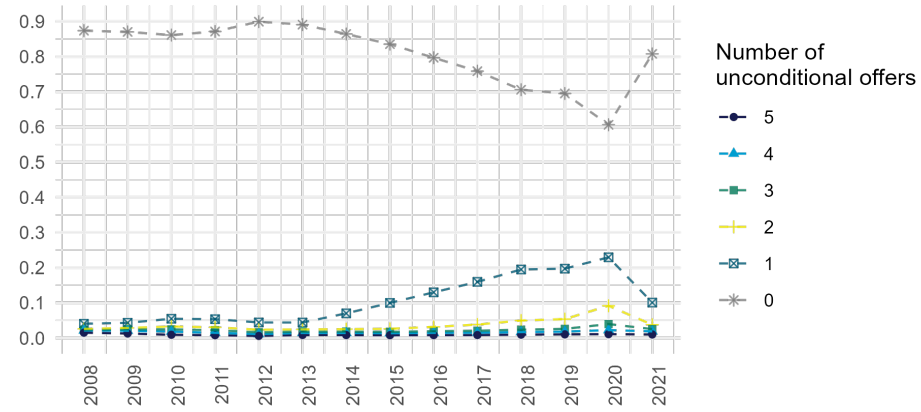
Notes: Restricted to having sent five applications.

Figure A3: Number of offers (as shares)

A) Conditional offers



B) Unconditional offers



Notes: Panels A and B show time trends in conditional (panel A) and unconditional offers (panel B), conditional on applicants having submitted five applications. All values are presented as shares, with categories illustrated in different colors. For total offers, please see Figure 4.

Table A5: Descriptive statistics of applicants by sequence of decisions

	Field defined as broad field (19 fields)				
	One field	One field at one HEI	One HEI	Mixed	Overall
Share female	0.56	0.49	0.45	0.55	0.56
Polar4	3.52	3.34	3.40	3.53	3.53
IMD	3.32	2.97	2.95	3.28	3.30
Parental SES:					
Low SES	0.17	0.24	0.22	0.16	0.17
Medium SES	0.44	0.45	0.48	0.45	0.45
High SES	0.23	0.13	0.12	0.22	0.22
Unknown	0.17	0.18	0.18	0.17	0.17
N	2,187,484	6,439	26,018	1,921,234	4,141,175
Shares	0.53	0.00	0.01	0.46	1.00

Note: Mean values pooled over the years 2008 to 2021. Field defined first digit of JACS code (19 different fields). Combined degrees are considered as separate fields. For definition of Polar4, IMD, and Parental SES, please refer to Table A1.

Table A6: Shares of field-institution combinations

	Number of fields based on:									
	Broad field (19 fields)					Narrow field (192 fields)				
	1	2	3	4	5	1	2	3	4	5
No. of Inst.										
1	0.002	0.003	0.002	0.001	0.000	0.000	0.001	0.002	0.003	0.002
2	0.007	0.012	0.009	0.003	0.000	0.002	0.006	0.010	0.009	0.004
3	0.020	0.026	0.015	0.004	0.000	0.008	0.018	0.022	0.013	0.004
4	0.056	0.067	0.026	0.004	0.000	0.027	0.060	0.043	0.019	0.004
5	0.445	0.222	0.068	0.007	0.000	0.324	0.238	0.130	0.044	0.007

Note: Share of different field-institution combinations. Mean values pooled over the years 2008 to 2021. The colors highlight the groups as in Panel C of Figure 4. The non-coloured cells are the "mixed"-category.

Table A7: Descriptive statistics of applicants by number of applications within home region

	Number of applications within home region						
	0	1	2	3	4	5	Overall
Share female	0.55	0.54	0.55	0.57	0.59	0.58	0.56
Polar4	3.63	3.60	3.44	3.32	3.31	3.54	3.53
IMD	3.52	3.51	3.29	3.05	2.78	2.97	3.30
Parental SES:							
Low SES	0.13	0.14	0.17	0.21	0.23	0.21	0.17
Medium SES	0.44	0.45	0.45	0.45	0.43	0.44	0.45
High SES	0.27	0.25	0.21	0.17	0.14	0.15	0.22
Unknown	0.16	0.16	0.16	0.18	0.20	0.19	0.17
Total N	1,281,551	902,402	646,964	422,659	287,671	599,928	4,141,175
Share	0.31	0.22	0.16	0.10	0.07	0.14	1.00

Note: Mean values pooled over the years 2008 to 2021. For definition of Polar4, IMD, and Parental SES, please refer to table A1. Regions are: Scotland, Northern Ireland, Wales as well as for England the regions East Midlands, East of England, London, North East, North West, South East, South West, West Midlands, and Yorkshire and the Humber.

Table A8: Descriptive statistics of applicants by number of offers received

	Number of offers an applicant received						
	0	1	2	3	4	5	Overall
Share female	0.74	0.66	0.60	0.54	0.52	0.54	0.56
Polar4	3.28	3.38	3.46	3.54	3.58	3.55	3.53
IMD	2.94	3.08	3.18	3.27	3.34	3.39	3.30
Parental SES:							
Low SES	0.23	0.2	0.19	0.17	0.16	0.15	0.17
Medium SES	0.43	0.44	0.44	0.44	0.44	0.45	0.45
High SES	0.13	0.17	0.19	0.22	0.24	0.24	0.22
Unknown	0.21	0.19	0.18	0.17	0.17	0.16	0.17
N	138,761	277,395	431,950	668,853	1,093,141	1,531,075	4,141,175
Shares	0.03	0.07	0.10	0.16	0.26	0.37	1.00

Note: Mean values pooled over the years 2008 to 2021. Offers include conditional and unconditional offers as of the June-30-deadline. For definition of Polar4, IMD, and Parental SES, please refer to Table A1.

B Joint Academic Coding System (JACS)

The Joint Academic Coding System (JACS) is a subject coding system for higher education in the UK, jointly owned by UCAS and HESA (see Table B1). All subjects are grouped into 164 two digit principal subjects.

Table B1: Joint Academic Coding System JACS 3.0

JACS subject area	**JACS principal subject**
Medicine & dentistry	(A0) Broadly-based programmes within medicine & dentistry
	(A1) Pre-clinical medicine
	(A2) Pre-clinical dentistry
	(A3) Clinical medicine
	(A4) Clinical dentistry
	(A9) Others in medicine & dentistry
Subjects allied to medicine	(B0) Broadly-based programmes within subjects allied to medicine
	(B1) Anatomy, physiology & pathology
	(B2) Pharmacology, toxicology & pharmacy
	(B3) Complementary medicines, therapies & well-being
	(B4) Nutrition
	(B5) Ophthalmics
	(B6) Aural & oral sciences
	(B7) Nursing
	(B8) Medical technology
	(B9) Others in subjects allied to medicine
Biological sciences	(C0) Broadly-based programmes within biological sciences
	(C1) Biology
	(C2) Botany
	(C3) Zoology
	(C4) Genetics
	(C5) Microbiology
	(C6) Sport & exercise science
	(C7) Molecular biology, biophysics & biochemistry
	(C8) Psychology
	(C9) Others in Biological Sciences
Veterinary science	(D1) Pre-clinical veterinary medicine
	(D2) Clinical veterinary medicine & dentistry
Agriculture & related subjects	(D0) Broadly-based programmes within agriculture & related subjects
	(D3) Animal science
	(D4) Agriculture
	(D5) Forestry & arboriculture
	(D6) Food & beverage studies
	(D7) Agricultural sciences

Continued on next page

JACS subject area	JACS principal subject
	(D9) Others in veterinary sciences, agriculture & related subjects
Physical sciences	(F0) Broadly-based programmes within physical sciences
	(F1) Chemistry
	(F2) Materials science
	(F3) Physics
	(F4) Forensic & archaeological sciences
	(F5) Astronomy
	(F6) Geology
	(F7) Science of aquatic & terrestrial environments
	(F8) Physical geographical sciences
	(F9) Others in physical sciences
Mathematical sciences	(G1) Mathematics
	(G2) Operational research
	(G3) Statistics
	(G9) Others in mathematical sciences
Computer science	(I1) Computer science
	(I2) Information systems
	(I3) Software engineering
	(I4) Artificial intelligence
	(I5) Health informatics
	(I6) Games
	(I7) Computer generated visual & audio effects
	(I9) Others in Computer sciences
Engineering & technology	(H0) Broadly-based programmes within engineering & technology
	(H1) General engineering
	(H2) Civil engineering
	(H3) Mechanical engineering
	(H4) Aerospace engineering
	(H5) Naval architecture
	(H6) Electronic & electrical engineering
	(H7) Production & manufacturing engineering
	(H8) Chemical, process & energy engineering
	(H9) Others in engineering
	(J1) Minerals technology
	(J2) Metallurgy
	(J3) Ceramics & glass
	(J4) Polymers & textiles
	(J5) Materials technology not otherwise specified
	(J6) Maritime technology
	(J7) Biotechnology

Continued on next page

JACS subject area	JACS principal subject
	(J9) Others in technology
Architecture, building & planning	(K0) Broadly-based programmes within architecture, building & planning
	(K1) Architecture
	(K2) Building
	(K3) Landscape & garden design
	(K4) Planning (urban, rural & regional)
	(K9) Others in architecture, building & planning
Social studies	(L0) Broadly-based programmes within social studies
	(L1) Economics
	(L2) Politics
	(L3) Sociology
	(L4) Social policy
	(L5) Social work
	(L6) Anthropology
	(L7) Human & social geography
	(L8) Development studies
	(L9) Others in social studies
Law	(M0) Broadly-based programmes within law
	(M1) Law by area
	(M2) Law by topic
	(M9) Others in law
Business & administrative studies	(N0) Broadly-based programmes within business & administrative studies
	(N1) Business studies
	(N2) Management studies
	(N3) Finance
	(N4) Accounting
	(N5) Marketing
	(N6) Human resource management
	(N7) Office skills
	(N8) Hospitality, leisure, sport, tourism & transport
	(N9) Others in business & administrative studies
Mass communications & documentation	(P0) Broadly-based programmes within mass communications & documentation
	(P1) Information services
	(P2) Publicity studies
	(P3) Media studies
	(P4) Publishing
	(P5) Journalism
	(P9) Others in mass communications & documentation

Continued on next page

JACS subject area	JACS principal subject
Languages	(Q0) Broadly-based programmes within languages
	(Q1) Linguistics
	(Q2) Comparative literary studies
	(Q3) English studies
	(Q4) Ancient language studies
	(Q5) Celtic studies
	(Q6) Latin studies
	(Q7) Classical Greek studies
	(Q8) Classical studies
	(Q9) Others in linguistics, classics & related subjects
	(R1) French studies
	(R2) German studies
	(R3) Italian studies
	(R4) Spanish studies
	(R5) Portuguese studies
	(R6) Scandinavian studies
	(R7) Russian & East European studies
	(R8) European studies
	(R9) Others in European languages, literature & related subjects
	(T1) Chinese studies
	(T2) Japanese studies
	(T3) South Asian studies
	(T4) Other Asian studies
	(T5) African studies
	(T6) Modern Middle Eastern studies
	(T7) American studies
	(T8) Australasian studies
	(T9) Others in Eastern, Asiatic, African, American & Australasian languages, literature & related subjects
Historical & philosophical studies	(V0) Broadly-based programmes within historical & philosophical studies
	(V1) History by period
	(V2) History by area
	(V3) History by topic
	(V4) Archaeology
	(V5) Philosophy
	(V6) Theology & religious studies
	(V7) Heritage studies
	(V9) Others in historical & philosophical studies
Creative arts & design	(W0) Broadly-based programmes within creative arts & design
	(W1) Fine art
	(W2) Design studies
	(W3) Music

Continued on next page

JACS subject area	JACS principal subject
	(W4) Drama
	(W5) Dance
	(W6) Cinematics & photography
	(W7) Crafts
	(W8) Imaginative writing
	(W9) Others in creative arts & design
Education	(X0) Broadly-based programmes within education
	(X1) Training teachers
	(X2) Research & study skills in education
	(X3) Academic studies in education
	(X9) Others in education
Combined	(Y0) Combined

C Details on the survey

C.1 Survey Questions

In this section of the survey, we aim to gain a deeper understanding of your perceptions regarding students' university choice. The questions below refer to applications for undergraduate courses at UK higher education institutions sent via the Universities and Colleges Admissions Service (UCAS) between 2008 and 2021.

Question 1:
Applicants who apply for an undergraduate course via UCAS can send up to five applications via the so-called main scheme. What do you think is the average number of applications students sent in the period 2008 to 2021?
Open question, any value between 0 and 5 could be inserted (with a maximum of five decimal places).

Question 2:
UCAS distinguishes between 19 different subject areas (based on the first digit of Joint Academic Coding System/JACS 3.0). These are:

- Medicine and Dentistry
- Subjects allied to medicine
- Biological Sciences
- Veterinary Sciences
- Agriculture and related studies
- Physical Sciences
- Mathematical Sciences
- Computer Science
- Engineering and Technology
- Architecture, Building and Planning
- Social Studies
- Law
- Business and Administrative Studies
- Mass Communications and Documentation
- Languages
- Historical and Philosophical Studies
- Creative Arts and Design
- Education
- Combined Degrees

Considering only those applicants who submitted **five** applications, to how many different fields, as defined by the aforementioned JACS codes, do you think applicants applied to on average in the period 2008 to 2021?
Open question, any value between 0 and 5 could be inserted (with a maximum of five decimal places).

Question 3:
Considering only those applicants who submitted **five** applications, to how many different higher education institutions do you think applicants applied to on average in the period 2008 to 2021?
Open question, any value between 0 and 5 could be inserted (with a maximum of five decimal places).

Question 4:
Students can receive three different responses by an institution:

(i) they can be rejected

(ii) can be unconditionally accepted

(iii) or can receive an offer conditional on their secondary school results, the so-called A levels.

Considering only those applicants who submitted **five** applications, how many offers (conditional or unconditional) do you think the applicants received on average in the period 2008 to 2021?
Open question, any value between 0 and 5 could be inserted (with a maximum of five decimal places).

Question 5:
Please consider **female and male** applicants. Do you think that regarding their applications:

(Option 1) Female applicants send on average FEWER applications compared to male applicants.

(Option 2) Female and male applicants send on average around the SAME number of applications compared to male applicants.

(Option 3) Female applicants send on average MORE applications compared to male applicants.

(Option 4) I have no clue.

One out of the four possible options could be checked.

Question 6:
Please consider **female and male** applicants and only those applicants who submitted **five** applications. Do you think that regarding their offers:

(Option 1) Female applicants receive on average FEWER offers compared to male applicants.

(Option 2) Female and male applicants receive on average around the SAME number of offers compared to male applicants.

(Option 3) Female applicants receive on average MORE offers compared to male applicants.

(Option 4) I have no clue.

One out of the four possible options could be checked.

Question 7:
Please consider applicants with **high and low socio-economic background**, where for low socio-economic background (LOW SES) we intend applicants with the household head working in a routine or semi-routine occupation, while for high socio-economic background (HIGH SES) we intend applicants who have at least one parent working in higher managerial or professional occupations. Do you think that:

(Option 1) Applicants with LOW SES send on average FEWER applications compared to applicants with HIGH SES.

(Option 2) Applicants with LOW SES send on average around the SAME number of applications compared to applicants with HIGH SES.

(Option 3) Applicants with LOW SES send on average MORE applications compared to applicants with HIGH SES.

(Option 4) I have no clue.

One out of the four possible options could be checked.

Question 8:
Considering applicants with **high and low socio-economic background** as before and only those applicants who submitted **five** applications. Do you think that regarding offers:

(Option 1) Applicants with LOW SES receive on average FEWER offers compared to applicants with HIGH SES.

(Option 2) Applicants with LOW SES receive on average around the SAME number of offers compared to applicants with HIGH SES.

(Option 3) Applicants with LOW SES receive on average MORE offers compared to applicants with HIGH SES.

(Option 4) I have no clue.

One out of the four possible options could be checked.

In this section of the survey, we will focus on your current employment and past involvement in undergraduate admissions.

Question 9:
What is your current main employment or affiliation?

(Option 1) University or research institute in the UK

(Option 2) University or research institute in Europe excluding the UK

(Option 3) University or research institute in the US

(Option 4) University or research institute elsewhere

(Option 5) Other than university or research institute

One out of the five possible options could be checked.

Question 10:
What is your main discipline?

(Option 1) Economics

(Option 2) Education

(Option 3) Sociology

(Option 4) Other social sciences

(Option 5) Sciences, Technology, Engineering, Mathematics

(Option 6) Other

One out of the six possible options could be checked.

Question 11:
Have you ever been involved with undergraduate students' admissions?

(Option 1) Yes, in a UK higher education institution

(Option 2) Yes, elsewhere

(Option 3) No

One out of the three possible options could be checked.

Question 12:
Please indicate your gender identity:

(Option 1) Male

(Option 2) Female

(Option 3) Non-binary/third gender/other

One out of the possible options could be checked.

C.2 E-Mail invitation

E-Mail to people within the field of Economics of Education

Sender: Uni Choice Survey
Subject: Experts' Views on University Choices

Dear Colleague,

We hope this message finds you well. We are writing to you as an academic with expertise in education.

We are conducting an anonymous survey **to understand academics' perceptions regarding undergraduate students' university choices**. This project involves Friederike Hertweck (RWI – Leibniz Institute for Economic Research, Germany), Mirco Tonin (Free University of Bozen-Bolzano, Italy), and Michael Vlassopoulos (University of Southampton, UK).

We would like to invite you to participate in a short survey, which should take no more than 10 minutes. Your participation will enhance the perspective on the topic and contribute to a more comprehensive understanding of transitions into higher education. As a token of appreciation, the research team will donate 1 GBP to the charity "Save the Children UK" for every completed survey (up to a maximum of 150 GBP).
To participate, please follow this link:
Take the Survey

Or copy and paste the URL below into your internet browser:
Link

We sincerely value your time and thoughtful contribution to our survey. If you have any inquiries or require further information, please feel free to reach out to *E-Mail* or by responding to this email.

Best regards,
Friederike Hertweck, Mirco Tonin, Michael Vlassopoulos

Follow the link to opt out of future emails:
Link

E-Mail to UK academics

Sender: Uni Choice Survey
Subject: Experts' Views on University Choices
Dear Colleague,

We hope this message finds you well. We are writing to you as an academic at a UK higher education institution.

We are conducting an anonymous survey **to understand academics' perceptions regarding undergraduate students' university choices**. This project involves Friederike Hertweck (RWI – Leibniz Institute for Economic Research, Germany), Mirco Tonin (Free University of Bozen-Bolzano, Italy), and Michael Vlassopoulos (University of Southampton, UK).

We would like to invite you to participate in a short survey, which should take no more than 10 minutes. Your participation will enhance the perspective on the topic and contribute to a more comprehensive understanding of transitions into higher education. As a token of appreciation, the research team will donate 1 GBP to the charity "Save the Children UK" for every completed survey (up to a maximum of 150 GBP).
To participate, please follow this link:
Take the Survey

Or copy and paste the URL below into your internet browser:
Link

We sincerely value your time and thoughtful contribution to our survey. If you have any inquiries or require further information, please feel free to reach out to *E-Mail* or by responding to this email.

Best regards,
Friederike Hertweck, Mirco Tonin, Michael Vlassopoulos

Follow the link to opt out of future emails:
Link

Survey: Starting/Introductory page

Study Title:
A Survey of Experts' Views on University Choices

Researchers:
Michael Vlassopoulos (University of Southampton)
Mirco Tonin (Free University of Bozen-Bolzano)
Friederike Hertweck (RWI – Leibniz Institute for Economic Research)

What is the research about?
We are conducting an anonymous survey to understand academic perceptions regarding undergraduate students' university choices. This study was approved by the Faculty Research Ethics Committee (FREC) at the University of Southampton (Ethics/ERGO Number: 93571).

What will happen to me if I take part?
This study involves completing an anonymous questionnaire which should take approximately 5 to 10 minutes of your time. If you are happy to complete this survey, you will need to tick (check) the box below to show your consent. As this survey is anonymous, we will not be able to know whether you have participated, or what answers you provided.

Why have I been asked to participate?
You have been asked to take part because we believe you have relevant expertise in Higher Education. We are aiming to recruit around 150 participants for this study.

What information will be collected?
The questions in this survey ask for information in relation to your perceptions of applications to higher education. The questions also ask for some general information related to you and your area of research. You do not have to answer all the questions if you do not wish to do so.

What are the possible benefits of taking part?
If you decide to take part in this study, you will not receive any direct benefits; however, as a token of appreciation, we will donate 1 GBP to the charity "Save the Children UK" for every completed survey (up to a maximum of 150 GBP). Your participation will also enhance the perspective on the topic and contribute to a more comprehensive understanding of transitions into higher education.

Are there any risks involved?
It is expected that taking part in this study will not cause you any psychological discomfort and/or distress, however, should you feel uncomfortable you can leave the survey at any time.

What will happen to the information collected?
All information collected for this study will be stored securely on a password protected computer and backed up on a secure server. In addition, all data will be pooled and only compiled into data summaries or summary reports. Only the research team will have access to this information. The information collected will be analysed and presented on conferences and published in a journal. The University of Southampton conducts research to the highest standards of ethics and research integrity. In accordance with our Research Data Management Policy, data will be held for 10 years after the study has finished when it will be securely destroyed.

What happens if there is a problem?
If you are unhappy about any aspect of this study and would like to make a formal complaint, you can contact the Head of Research Integrity and Governance, University of Southampton, on the following contact details: Email: *E-mail and phone*Please quote the Ethics/ERGO number above. Please note that by making a complaint you might be no longer anonymous.More information on your rights as a study participant is available via this link. Thank you for reading this information sheet and considering taking part in this research.